N73 24693

PHYSICS IN PERSPECTIVE

VOLUME II
PART B

The Interfaces

NATIONAL
ACADEMY of
SCIENCES

PHYSICS IN PERSPECTIVE

VOLUME II
PART B

The Interfaces

Physics Survey Committee • National Research Council

NATIONAL ACADEMY OF SCIENCES Washington, D.C. 1973

NOTICE: The study reported herein was undertaken under the aegis of the Committee on Science and Public Policy (COSPUP) of the National Academy of Sciences–National Research Council, with the express approval of the Governing Board of the National Research Council.

Responsibility for all aspects of this report rests with the Physics Survey Committee, to whom sincere appreciation is here expressed.

The report has not been submitted for approval to the Academy membership or to the Council but, in accordance with Academy procedures, has been reviewed and approved by the Committee on Science and Public Policy. It is being distributed, following this review, with the approval of the President of the Academy.

LIBRARY OF CONGRESS CATALOGING IN PUBLICATION DATA

National Research Council. Physics Survey Committee.
Physics in perspective.

"Undertaken under the aegis of the Committee on Science and Public Policy of the National Academy of Sciences–National Research Council."
Includes bibliographical references.
PARTIAL CONTENTS: v. 2. pt. A. The core subfields of physics.–v. 2. pt. B. The interfaces.
1. Physics–Research–United States. 2. Science–United States. I. National Research Council. Committee on Science and Public Policy. II. Title.
QC9.U5N28 530'.0973 72-84752
ISBN 0-309-02101-4

Available from
Printing and Publishing Office, National Academy of Sciences
2101 Constitution Avenue, Washington, D.C. 20418

Printed in the United States of America

Committee on Science and Public Policy

MELVIN CALVIN, University of California, Berkeley, *Chairman*
HARVEY BROOKS, Harvard University, *Past Chairman*
GEORGE BACKUS, University of California, San Diego
GERALD CLEMENCE, Yale University
PAUL DOTY, Harvard University
ROBERT GARRELS, University of Hawaii
ARTHUR HASLER, University of Wisconsin
LELAND J. HAWORTH, Associated Universities, Inc.
RUDOLF KOMPFNER, Bell Telephone Laboratories, Inc.
JAY LUSH, Iowa State University
ROBERT K. MERTON, Columbia University
HERSCHEL L. ROMAN, University of Washington
HARRISON SHULL, Indiana University
I. M. SINGER, Massachusetts Institute of Technology

ROBERT E. GREEN, National Academy of Sciences
Executive Secretary

Physics Survey Committee

D. ALLAN BROMLEY, Yale University, *Chairman*
*DANIEL ALPERT, University of Illinois
RAYMOND BOWERS, Cornell University
JOSEPH W. CHAMBERLAIN, The Lunar Science Institute
HERMAN FESHBACH, Massachusetts Institute of Technology
GEORGE B. FIELD, University of California, Berkeley
RONALD GEBALLE, University of Washington
CONYERS HERRING, Bell Telephone Laboratories, Inc.
ARTHUR R. KANTROWITZ, AVCO-Everett Research Laboratory
THOMAS LAURITSEN, California Institute of Technology
FRANKLIN A. LONG, Cornell University
E. R. PIORE, International Business Machines Corporation
E. M. PURCELL, Harvard University
ROBERT G. SACHS, University of Chicago
*CHARLES H. TOWNES, University of California, Berkeley
ALVIN M. WEINBERG, Oak Ridge National Laboratory
VICTOR F. WEISSKOPF, Massachusetts Institute of Technology

*Members of original Survey Committee. Other responsibilities precluded these physicists from participating in other than the early deliberations.

Ex officio

HARVEY BROOKS, Harvard University
E. L. GOLDWASSER, National Accelerator Laboratory
H. WILLIAM KOCH, American Institute of Physics
ROMAN SMOLUCHOWSKI, Princeton University

CONSULTANTS

MARVIN GOLDBERGER, Princeton University
ROBERT W. KEYES, International Business Machines Corporation

LIAISON REPRESENTATIVES*

C. NEIL AMMERMAN, Department of Defense
W. O. ARMSTRONG, National Aeronautics and Space Administration
MARCEL BARDON, National Science Foundation
DONALD E. CUNNINGHAM, National Science Foundation
PAUL F. DONOVAN, National Science Foundation
WAYNE R. GRUNER, National Science Foundation
GEORGE A. KOLSTAD, Atomic Energy Commission
JOHN J. McCAMBRIDGE, Department of Defense
WILLIAM METSCHER, Department of Defense
SIDNEY PASSMAN, National Science Foundation
JACK T. SANDERSON, National Science Foundation
A. W. SCHARDT, National Aeronautics and Space Administration
MILTON M. SLAWSKY, Department of Defense
HENRY SMITH, National Aeronautics and Space Administration

GEORGE W. WOOD, National Academy of Sciences, *Executive Secretary*

Staff Officers

BERTITA E. COMPTON
BRUCE N. GREGORY
CHARLES K. REED

*In general, no agency was represented on the Committee or any of its Panels by more than one person at a time.

Preface

The Physics Survey Committee was appointed by the President of the National Academy of Sciences in mid-1969 to survey the status, opportunities, and problems of physics in the United States. Volume I of *Physics in Perspective* constitutes the full report of the Committee. This, and the companion parts of Volume II, complete the report of the survey.

The Survey Committee early concluded that it was essential that it obtain detailed information from experts in each of a number of physics subfields and interface areas. For each of these subfields and interface areas a chairman was appointed by the Chairman of the Survey Committee, and groups of recognized experts were brought together to survey and report on their respective subject areas.

Several of the subfields have relatively well-defined and traditional boundaries in physics. Included are the core subfields of acoustics, optics, condensed matter, plasmas and fluids, atomic, molecular, and electron physics, nuclear physics and elementary-particle physics. The reports of these panels constitute Part A of Volume II. In addition, there are several important interface areas between physics and other sciences. In the case of astronomy, where activity is particularly vigorous at the interface and is overlapping, the Physics and Astronomy Survey Committees agreed to form a joint panel that would report on astrophysics and relativity, an area of special interest to both. The broad area in which physics overlaps geology, oceanography, terrestrial and planetary atmospheric studies, and other environmental sciences was defined as earth and planetary physics, and a panel was established to survey it. In covering the physics–chemistry and physics–biology interfaces, the broader designations "physics in chemistry" and "physics in biology" were chosen to avoid restricting the work of the panels to the already traditional boundaries of these interdisciplinary fields.

Although each panel—and particularly those responsible for the core subfields—was asked to consider the interaction of its subfield with technology, the Committee anticipated that the emphasis would be on recent developments that advanced the state of the art and on what is generally described as *high technology*. Therefore, to include more specifically the active instrumentation interface between physics and the more traditional manufacturing sectors of the economy—steel, drugs, chemicals and consumer goods, to name only a few, in which many old parameters are being measured and controlled in new and ingenious ways—a separate panel was established.

Panels were also appointed to centralize the statistical data-collection activities of the survey and to address the questions of physics in education and education in physics. Each of these panels prepared a report, and, in addition, an extended report on the dissemination and use of the information of physics was prepared by a member of the Committee. With the exception of the one on statistical data, all these reports are included in Part B of Volume II; the Statistical Data Panel report constitutes Part C of Volume II.

The Nuclear Physics Panel was commissioned to carry out its survey on an accelerated time scale, and in greater depth than the other panels, in response to a specific request from the President's Science Advisory Committee (PSAC) for findings and recommendations that could be used in policy and planning discussions at an early date. The final report of that panel, which appears in Part A, is a revised and updated version of the one transmitted to PSAC in 1971.

A number of subjects in classical physics, such as mechanics, heat, thermodynamics, and some elements of statistical physics, were not considered explicitly in the survey. This omission is in no sense intended to imply any lack of importance of these fields but merely indicates that they are mature fields in which relatively little research *per se* is currently being conducted.

In the very nature of the survey, the Committee and its panels have explored many alternatives and options in developing their reports. It should thus be emphasized that the lack of explicit mention of any one of these does not imply that it has not been considered or examined.

Early in the survey, the Committee developed and addressed to each panel a lengthy charge, which appears as Appendix A. This charge was broad-ranging and dealt with the structure and activity of a subfield, viewed not only internally but also in terms of its past, present, and potential contributions to other physics subfields, other sciences, technology, and society generally. Consonant with the overall survey objectives, each panel was asked to develop several detailed budgetary projections ranging from one that would permit exploitation of all currently identified opportunities in a subfield to one that continued to decrease during the period under consideration.

Clearly, the charge was most directly relevant to the more traditional subfields; in the case of the interface panels, some questions were inevitably unanswerable without a survey of equivalent scope of the field or fields on the other side of the interface. In astronomy, such a survey was available. Nevertheless, from the reports included in this and in the companion parts of Volume II, it is plain that the panels have responded in depth to the questions asked.

Initial draft responses to the charge were presented to the Survey Committee by the panel chairmen during an extended working session in June 1970, and, following subsequent discussions and reviews, preliminary panel reports were submitted to the Committee during the summer of 1971. Whenever possible, each of these preliminary reports was forwarded for comment to a group of some ten readers, selected jointly in each case by the panel chairman and the chairman of the appropriate division of the American Physical Society or other Member Society of the American Institute of Physics. These readers were chosen, insofar as possible, from among the most active scientists in each subfield, with particular emphasis on younger scientists who had not been involved previously in the survey. The Committee received excellent cooperation from all of them. They provided fresh insight and new viewpoints on many aspects of the panel reports. Their comments and those of the Survey Committee and other reviewers were carefully considered by the panels in the preparation of the final reports that appear herein.

It must be emphasized that the panel reports and their recommendations and conclusions were addressed specifically to the Survey Committee. The many instances in which the Committee concurred with and supported these findings are reflected in the Committee's report, Volume I. On occasion, however, the Committee, from its broader viewpoint covering not only all of physics but also its broader external interaction, not unexpectedly reached somewhat different conclusions.

The panel reports are being made available here in the form submitted to the Survey Committee, not only to provide the detailed technical background and documentation for many of the Committee's findings, but also because they provide, to a unique degree, a measure of the vitality and strength of the different subfields of physics. Repeatedly in its activity, the Survey Committee has been reminded of the unity of physics and, indeed, of all science. This intellectual thread is interwoven through all the panel reports.

The Survey Committee is profoundly grateful to the members of its panels and most particularly to their chairmen, for their effective and thoughtful responses to the often difficult questions posed to them. Perhaps the most difficult have been those relating to the future style, direction, and thrust of physics under conditions in which not even all those projects and groups judged excellent by peer and support agency reviews can hope to find support. These questions are much more directly answerable in some subfields than in others—in those dependent upon very large facilities rather than on more modest instrumental requirements—but they are very significant in all subfields.

The panel chairmen responded frequently and effectively to Committee requests for additional information and assistance; they participated fully in a number of the major Committee working sessions and they gave most generously of their time and effort throughout the survey.

Support for the survey activity has been provided equally by the Atomic Energy Commission, the Department of Defense, the National Aeronautics and Space Administration, and the National Science Foundation. Additional assistance has been provided through grants from the American Physical Society and from the American Institute of Physics.

Staff of all the federal agencies engaged in the support of physics have given generously of their time and effort in searching out and providing answers to innumerable questions. Liaison representatives of these agencies participated in many long days of discussion as the Committee and panel reports developed. The Committee is deeply grateful to all of them.

The Committee and its panels cannot hope to acknowledge in detail all the assistance that they have received from many persons and organizations throughout the country. Over and above their major contributions to the activity of the Survey Committee itself, George W. Wood, Charles K. Reed, Bruce N. Gregory, and Bertita E. Compton have worked directly with the different panels in many, many ways and have provided an overall coherence that otherwise would have been quite impossible. They deserve our particular gratitude. Jacqueline Boraks has accomplished the often overwhelming task of copy editing the entire survey report with remarkable effectiveness and taste. In these panel reports she has been ably assisted by Jeannette W. Lindsay. And finally, Beatrice Bretzfield, the Secretary to the Physics Survey at the Academy, and Mary Anne Thomson, my administrative assistant at Yale, have been of tremendous help to the Committee and to me throughout the survey. To all of these, I would express both my personal thanks and that of the Committee and its panels.

D. ALLAN BROMLEY, *Chairman*
Physics Survey Committee

Contents

VIII
Astrophysics and Relativity

PANEL MEMBERS*

GEORGE B. FIELD, University of California, Berkeley, *Chairman*
GEOFFREY BURBIDGE, University of California, San Diego
GEORGE W. CLARK, Massachusetts Institute of Technology
DONALD D. CLAYTON, Rice University
ROBERT H. DICKE, Princeton University
KENNETH KELLERMANN, National Radio Astronomy Observatory
CHARLES MISNER, University of Maryland
EUGENE N. PARKER, University of Chicago
EDWIN E. SALPETER, Cornell University
MAARTEN SCHMIDT, California Institute of Technology
STEVEN WEINBERG, Massachusetts Institute of Technology

CONSULTANT

DAVID D. CUDABACK, University of California, Berkeley

LIAISON REPRESENTATIVES†

NANCY ROMAN, National Aeronautics and Space Administration
A. W. SCHARDT, National Aeronautics and Space Administration
HAROLD S. ZAPOLSKY, National Science Foundation

READERS

ELIHU A. BOLDT, National Aeronautics and Space Administration
GIOVANNI G. FAZIO, Smithsonian Institution Astrophysical Observatory
HUGH M. JOHNSON, Lockheed Palo Alto Research Laboratory
D. P. McNUTT, Naval Research Laboratory
ALAN T. MOFFET, California Institute of Technology
J. A. SIMPSON, University of Chicago
JAMES W. TRURAN, Yeshiva University

* Joint Panel with the Astronomy Survey.
† In general, no agency was represented on the panel by more than one person at a given time.

Preface

The Panel was established jointly in 1969 by the Physics and the Astronomy Survey Committees. Its charge was to review the scientific status of and the research methods used in astrophysics and relativity and, in the light of opportunities for future advance, to recommend a program that would exploit the opportunities at various levels of available funding. The present report is the response to that charge.

The parent committees are concerned with the health of the entire national effort in physics and in astronomy, of which astrophysics and relativity is only a small part. Because many of the facilities recommended in this report have applications outside of the subfield (particularly in other branches of astronomy), the parent committees had to reconcile the recommendations of this report with those of several other panels. This report should therefore be read in the context of the reports of both the Physics and Astronomy Survey Committees.

The nonspecialist may find parts of the report too technical. In that event, maximum benefit will be derived from reading Chapters 1 (Introduction), 2 (Summary and Recommendations), 3 (The Impact of Cosmology on Culture and Science), and 8 (Manpower, Funding, and Education).

Contents

VIII
ASTROPHYSICS AND RELATIVITY

1 The Nature of the Field and Scope of the Report

How should we define the nature and scope of astrophysics and relativity? This Panel interprets it as the study of those phenomena in the astronomical universe that require Einstein's theory of general relativity for their explanation. For example, a star composed of neutrons has such strong gravitational fields that the Newtonian theory of gravitation breaks down and the theory of general relativity is needed.

The application of relativity to astronomical phenomena continues a tradition established earlier when other branches of physics, such as Newtonian mechanics, atomic and nuclear physics, and magnetohydrodynamics, were so applied. Although our title might indicate that we are concerned with all such applications (that is, all astrophysics), for the sake of brevity we have limited our discussion to only those branches of astrophysics that are intimately connected with relativistic problems. Other branches are covered in other panel reports, particularly of the Astronomy Survey.

Although the scientists engaged in astrophysics and relativity are principally physicists and astronomers, this subfield includes the whole of cosmology—the study of the origin and evolution of the universe—a subject that interests all educated men. Not only does cosmology play a unifying role with respect to other evolutionary sciences, such as biology and geology, but it touches on fundamental questions of broader interest, such as the origin of matter and the nature of time, that are rooted in ancient religious and philosophical traditions. We have treated cosmology as the central theme of our report, because, in exemplifying the connections with other human endeavors, it may be helpful to the nonscientist who wishes to assess the value of scientific effort in astrophysics and relativity.

In Chapter 2 the reader will find a summary of the report, together with our recommendations to the Physics Survey Committee. After discussing the cultural impact of the field in Chapter 3, we show in Chapter 4 how traditional relativistic cosmology has been revitalized by recent—and often unexpected—astronomical discoveries. Then in Chapter 5 we sketch in some detail the main lines of advance at the present time, indicating those in which additional effort would be most rewarding. The relation to other branches of science is considered in Chapter 6. Chapter 7 discusses the testing of general relativity. In Chapter 8 we describe and evaluate the resources in manpower and funding that are needed to exploit research opportunities in astrophysics and relativity to the fullest.

Because the theory of gravitation is properly a part of physics and because the conditions under which general relativity is operative can occur only on an astronomical scale, it is appropriate that this report be submitted to both the physics and astronomy communities, as well as to the government and other agencies that support them, through the Physics Survey Committee and the Astronomy Survey Committee of the National Research Council. These committees are reviewing the whole of physics and astronomy, of which astrophysics and relativity is only a small part.

2 Summary and Recommendations

Astrophysics and relativity is another name for the current attempt by scientists to understand the basic overall structure of the physical world in which we live. The current effort, begun early in the twentieth century, is based on the ideas of space and time conceived by Albert Einstein. If it is successful, it may be comparable in its impact with the world views established by the Greeks of antiquity and by Newton and his followers in the eighteenth century. When we remember that the Greek world view was the seed of what we now think of as Western civilization and that Newtonian physics was the scientific basis for the Industrial Revolution, we begin to understand why the triumph of Einstein's ideas about the world could have far-reaching impact.

Einstein had the notion that ordinary matter, if collected in sufficient quantity, can make space in its vicinity "curved." He developed this idea into the theory of general relativity, which treats the interaction of matter with matter on a large scale, through gravitation. Physicists working with this theory predicted that the universe of stars and galaxies should be either expanding or contracting and that the galaxies should be moving with tremendous speeds. In the 1920's, astronomers studying distant galaxies found that they were doing exactly what Einstein's theory predicts, thereby opening a new era of study of the effects of general relativity. The spectrum of each galaxy is shifted toward longer wavelengths (red shift), indicating a velocity away from us that is proportional to its distance. Although it was found later that Newtonian models also expand or contract, they are not capable of explaining the behavior of galaxies at distances so great that the velocity–distance relation would predict velocities of about the speed of light.

Einstein noted that the effects of general relativity should be present in the solar system. Although they are small and difficult to detect, all three effects that he predicted have been observed to have approximately the Einstein value. Recently a rival theory of relativity (the scalar–tensor theory) has been developed that predicts somewhat different values for relativistic effects in the solar system. The present precision of tests within the solar system is not sufficient to choose unambiguously between the two theories, but improvements expected in the near future should make this choice possible.

2.1 THE EXPANDING UNIVERSE AND COSMOLOGY

Since the discovery of the expansion of the universe, cosmology (the study of the structure and the evolution of the universe) has been based on relativistic models of space and time. The Friedmann models, based on Einstein's theory, postulate that the universe is uniform on the large scale and that its expansion began with a huge explosion several billion years ago. Then, according to Einstein's equations, it continued to expand, although it was constantly decelerated by the retarding effects of gravitational attraction. Two parameters are required to describe completely the behavior of the universe for all time: H_0, the Hubble constant, which is the rate of increase of expansion velocity with distance, and q_0, the deceleration parameter, which tells whether the average density of matter in the universe is sufficiently large that its gravitational deceleration can ultimately halt the expansion and cause the universe to implode upon itself. The scalar–tensor theory, which leads to qualitatively similar results, is specified by an additional parameter, ω, the coupling constant for the scalar field, which can be determined by observations in the solar system.

A rival cosmological theory, the steady-state model, has been derived from a modification of Einstein's equations that permits the continuous creation of new matter throughout the universe. The new matter replaces that lost by expansion, and the universe is in a steady state. Continuous creation is contrary to the present quantum theory of elementary particles, but the rate of creation required is so low (one atom per cubic foot in ten billion years) that it has not been ruled out experimentally.

Cosmology is intimately related to other sciences. If the steady-state model is correct, our understanding of particle physics must be modified; if the big-bang model is correct, the extremely high density and corresponding high temperature of the initial explosion ("fireball") will require extension of the theory of elementary particles to extremely high particle energies. For example, quarks, the hypothetical components of elementary particles, might have been present in the big bang and might even have survived in small numbers to the present era. Previous experience with the application of physics to astronomy, from Newton's explanation of planetary orbits to Bethe's explanation of the energy source of stars, suggests that a deep understanding of the big bang would be accompanied by a corresponding advance in physics.

Obviously all other parts of astronomy are strongly affected by cosmology. For example, in the big-bang theory galaxies cannot form from the hot material of the fireball itself, so the galaxies that we see now must have formed later, after the matter cooled, perhaps 100 million years after the big bang. The helium found in galaxies is partially the result of nuclear reactions in the big bang; heavier elements are the result of later reactions within stars. Therefore, all stars should have helium, but the oldest stars should have no heavier elements. On the other hand, the scalar–tensor theory predicts that less helium is formed in the fireball than do the Friedmann models. In the steady-state model, there is no big explosion, so even the helium can be formed only in stars. In this case it should be absent from the oldest stars. Thus, the evolution of stars and galaxies, which forms the central theme of astronomy, can be interpreted only within a cosmological framework. The study of cosmology also has an indirect effect on observational astronomy: since the observable differences between models are greatest at large distances, the most interesting objects are invariably faint and can be studied only with large telescopes. Telescopes constructed for cosmological research can be used also for detailed studies of closer, brighter objects that are the backbone of astronomical research.

The evolutionary sciences of geology and biology, devoted to the unraveling of the origin and evolution of the planets and life, are also ultimately understood in a cosmological framework. For example, radioactivity present in the young galaxy became trapped in the planets as they formed and determined their temperatures and rate of solidification. The amount of such radioactivity depends on the age and prior history of the galaxy, thus on cosmology.

In principle it is possible to discover by observation whether the big-bang

(Friedmann or scalar–tensor version) or steady-state model is correct. Hubble initiated a program with the 100-in. telescope on Mt. Wilson to obtain the accurate distances and velocities of galaxies, hence the value of H_0, and, using the apparent brightness of faint galaxies as an indicator of their distance, he sought deviations from a linear velocity–distance relation that would indicate the value of q_0. By patient work with the 200-in. telescope on Palomar Mountain, Hubble's successors have found that $1/H_0$ is about ten billion years, so that an expansion time scale of that order is indicated. In parallel work on clusters of stars within the galaxy, it has been shown that the oldest stars have approximately the same age, suggesting that the galaxy formed soon after the big bang, as required by theory. Young stars have helium in their atmospheres but the evidence for helium in old stars is still ambiguous, so the big-bang model is not wholly confirmed as yet.

The determination of q_0 by optical observations of galaxies is not yet conclusive for two reasons. Even the 200-in. telescope has difficulty measuring galaxies sufficiently far away that the effects of different values of q_0 are large. Moreover, if we do live in a big-bang universe, galaxies are expected to evolve, so that those at great distances, seen as they were billions of years ago, might well have systematically different intrinsic brightnesses. In that case, their apparent brightness would not be a reliable guide to their distance. A substantial effort is directed toward obtaining a detailed understanding of stellar evolution. If such an understanding is achieved, it will aid in predicting galactic evolution and in resolving the problem of the distance of faint galaxies.

2.2 RADIO ASTRONOMY AND EXPLODING OBJECTS

After World War II, physicists and engineers in Great Britain and Australia who had been engaged in the development of radar turned their attention to the problem of detecting faint radio waves of natural origin from space. They confirmed a prewar discovery made in the United States that the galaxy emits radio waves in the frequency range between 10 and 1000 MHz. Study of the spectrum and polarization of the radiation showed that it comes from relativistic electrons moving through interstellar space at speeds approaching that of light, with energies of nearly 5 GeV. These electrons are curved into circular orbits by a weak magnetic field (a few millionths of that of the earth) and emit radiation at very high harmonics of the frequency with which they are gyrating in the magnetic field. Practically all strong cosmic radio sources emit by this "synchrotron mechanism," which takes its name from the earth-bound accelerators in which the same phenomenon is observed.

As radio astronomers succeeded in increasing the angular resolution of their telescopes, discrete radio sources were found. The two brightest, Cassiopeia A and Cygnus A, turned out to be, respectively, the remains of an exploding star (a supernova) and an energetic extragalactic object. The supernova phenomenon had already been known to optical astronomers as a brief, powerful flare ten billion times brighter than an ordinary star. The most famous of these, the supernova of A.D. 1054, was observed by court astronomers in ancient China; today we know the remains of it as a bright nebulosity in the constellation of Taurus–the Crab nebula. This nebula apparently consists of the envelope of a star that was explosively ejected at a speed of 1000 miles per second. The inner part of the Crab nebula visibly glows in a way that was not understood until radio astronomers found that it, too, is a radio source (the third brightest); we now know that both radio waves and visible light are synchrotron emission by fast electrons. Apparently supernovae can accelerate large numbers of particles (constituting about one millionth of the mass of the star) to GeV energies.

When Cygnus A, the second brightest radio source, was observed optically with the 200-in. telescope, it turned out to be an exploding galaxy about half a billion light-years from the earth. Its great brightness, in spite of its distance, demonstrates its enormous radio power, which is emitted by clouds of relativistic particles apparently escaping from the parent galaxy. The total number of relativistic particles in Cygnus A is over a million times the number of particles in the sun. Such energetic explosions are unprecedented and pose a serious problem for the theorist. The problem is exacerbated by the discovery of quasi-stellar radio sources, or quasars–objects that at radio wavelengths seem to be similar to radio galaxies but that look more like stars when studied optically.

Radio galaxies are similar in size and optical brightness to ordinary bright galaxies and conform to the cosmological relation between recession velocity and distance. Quasars are much smaller than ordinary galaxies. The velocities of those studied so far, found from their red shifts, fall in the range between 10 percent and 87 percent of the speed of light. If their velocities are due to the expansion of the universe, the quasars are at enormous distances and must have intrinsic brightnesses as much as 1000 times greater than those of ordinary bright galaxies in spite of their small sizes. This conclusion is all the more remarkable because some of them vary in brightness in periods of a month or less, indicating that an energy source vastly more powerful than a galaxy is compacted into a region only a tenth of a light-year across, compared with dimensions of 100,000 light-years for normal galaxies. This situation is so difficult to explain that some scientists are not convinced that the red shifts are cosmological in nature, but other explanations that have been proposed for the red shifts raise equally grave problems.

Because the radio galaxies are optically bright and conform to the Hubble relation, they are particularly suitable objects for studies designed to establish a better value for the cosmological deceleration parameter q_0. Present studies, extending to five billion light years, yield a value of q_0 that is contrary to the steady-state model and favors a big-bang model in which the universe will gradually slow its expansion and ultimately collapse. Counts of faint radio sources have been used to argue against the steady-state. However, these results are open to debate in view of various uncertainties.

Quasars would be even more exciting as cosmological probes, as some may be extremely distant, but unfortunately they seem to have a wide variety of intrinsic brightnesses, making them of little use for this purpose. A pressing problem is to discover the source of relativistic particles in quasars. Recent investigations using the extraordinary angular resolution afforded by very-long-baseline radio interferometers (one ten millionth of a degree of arc) show that the particles are emitted in bursts of a solar mass or more from a region smaller than one tenth of a light year.

Recently scientists have found that at the center of the Crab nebula there is an object that emits pulses of electromagnetic waves 30 times per second. Probably this "pulsar" is actually a rapidly rotating neutron star—a star so compressed by its gravitation that its density is comparable with that of the atomic nucleus (5×10^{14} times that of water). Under these conditions, electrons are forced into protons to form neutrons and the material solidifies. General relativity is required to describe such dense objects, of which about 60 have been found so far. Apparently the Crab pulsar contains a magnetic field about a trillion times that of the earth, so that as it spins huge electrical fields are generated; these in turn accelerate particles to highly relativistic speeds. These are the particles that cause the Crab nebula to shine. It is quite possible that cosmic rays—the relativistic particles hitting the earth at all times—were accelerated by other neutron stars in a similar manner. We believe that neutron stars form when the interior of a massive star collapses as it exhausts its nuclear fuel. The energy released in the collapse is deposited in the envelope, causing the explosion that we detect as a supernova and leaving behind expanding remnants like the Crab nebula. The explosion is so violent that the envelope is heated to billions of degrees, with the result that nuclear reactions take place among the light nuclei present. Calculations indicate that such reactions may be responsible for the heavy elements found in interstellar space and in the young stars formed there. Thus gravitational collapse may be indirectly responsible for synthesis of many of the chemical elements. It is possible that the quasars are a similar phenomenon on a vastly larger scale—perhaps the central portion of a galaxy collapses to a high-density nucleus, which spins rapidly and accelerates particles in a similar manner. In this case, the basic energy source is gravitational and requires general relativity for its explanation.

In both cases—quasars and pulsars—we are confronted with large masses in small volumes, and it is precisely this type of configuration in which general relativity plays a major role. This theory states that if the volume becomes too small (so that the gravitational escape velocity approaches the speed of light), the whole configuration collapses to a mathematical point in a short time. Pulsars and quasars are kept from collapsing in this way by internal pressure or rotation. Other masses may not be so lucky and may have already collapsed. Perhaps there are many objects of this sort that we have not seen precisely because they are so dense that radiation cannot escape from them. Such "black holes" are predicted by theorists but so far have eluded detection.

Before leaving radio astronomy, we must mention the recent discovery of cosmic blackbody radiation. This radiation, which was predicted by Gamow to be a relic of the big-bang fireball, appears to have the spectrum of a thermal source at 2.7 K and comes equally from all directions. Here we may be looking at a "cosmic photosphere"—the surface of the fireball, like that of a star, where radiation leaves the hot gas. This surface is so far away and, therefore, is expanding away from us so rapidly that its radiation is shifted from visible wavelengths to the microwave region. Its detection has encouraged advocates of the big-bang model and permits one to calculate (under certain reasonable assumptions) that the production of helium in the fireball should be 27 percent if the Friedmann models are correct. This prediction is not yet confirmed. Recent observations at 1-mm wavelength (where the blackbody radiation should peak) indicate radiation intensities in excess of a blackbody spectrum, casting some doubt on the big-bang interpretation of this radiation.

2.3 ASTRONOMY USING OTHER WAVELENGTHS AND PARTICLES

Following the extension of the astronomical electromagnetic spectrum into the radio regime, recent effort has pushed into the x- and gamma-ray region (100 eV to 100 MeV) and the infrared region (1 to 100 μm). It has been found that 99 percent of the energy emitted from the Crab pulsar is in x rays above 1000 eV, and that pulses of visible radiation are emitted as well. Several extragalactic objects, including radio galaxies, a quasar, and clusters of galaxies, have been identified as x-ray sources. New x-ray telescopes in satellites may show that most radio sources are also x-ray sources, as we would expect from the theory of synchrotron emission. If so, x rays could be the main energy carrier; therefore, their study is crucial to a proper interpretation of radio sources. In addition, 100-MeV gamma rays also are emitted from the galaxy. Probably these are the result of collisions of cosmic-ray protons with interstellar atoms and so permit us to "see" relativistic protons in a way that complements the picture of relativistic electron distributions from radio astronomy. Cosmic-ray research continues to throw light on cosmological questions in a number of ways. Because the relativistic particles observed near the earth come from the same population that inhabits interstellar and perhaps intergalactic space, studies of cosmic rays permit us to infer something of the properties of those regions. For example, we find that the synchrotron emission observed in the radio range implies that the interstellar magnetic field averages about 10^{-5} of the earth's field, so that the amount of energy in the magnetic field and in relativistic particles is about equal. If the same is true in radio galaxies and quasars, we can infer the total energy in those objects. Very-high-energy particles observed near the earth cannot be trapped in the galaxy, suggesting that they must come to us through intergalactic space. Since they would generate rays by colliding with any gas that is present there, measurements of the gamma-ray intensity yield upper limits on the amount of intergalactic gas, a number of importance in cosmology.

Because cosmic rays are the only sample of distant matter that we can obtain, they are crucial in assessing the abundances of the elements in the energetic objects where they were born, such as supernovae, pulsars, or perhaps even radio galaxies. In this way we learn something of the history of these violent relativistic objects.

If the most energetic particles are born in extragalactic objects, they are subject to collisions with the photons of the cosmic blackbody background. Thus, the shape of the particle energy spectrum should tell us whether such photons are in fact present in intergalactic space, as required by the cosmological interpretation of the blackbody background.

Infrared studies of exploding galaxies show that, surprisingly, some of them emit copiously at 10-μm and longer wavelengths. In some cases, the major emission is in the infrared. Understanding of this phenomenon will aid in identifying the energy sources of these objects. Investigations of neutrinos and gravitational waves are just beginning. As these radiations are notoriously difficult to detect, the results are uncertain, but already there are surprises. The sun appears to emit less than half of the number of neutrinos expected from its thermonuclear processes, and the galactic nucleus could be emitting many times the gravitational radiation that one might expect. Obviously, confirmation of these exciting results is necessary.

2.4 SUMMARY OF PRESENT STATUS

Observations of the expanding universe suggest a giant relativistic explosion—the big bang. To discover whether the universe will continue to expand or ultimately collapse, and whether new matter arises to take the place of the old, we must have the correct theory of relativity and data about distant objects in the universe. Therefore, careful testing of relativity in the solar system and extensive optical, radio, x-ray, and infrared studies of distant objects using large telescopes will be necessary.

Intense explosions also occur on the scale of stars and galaxies. They are remarkably efficient in accelerating matter to relativistic speeds and are characterized by the concentration of matter in a small volume. This finding suggests the release of large amounts of gravitational energy—so large that the theory of general relativity is needed for the interpretation. Again, the correct theory of relativity is necessary, as well as studies at all wavelengths of the structures, spectra, and time variations of these objects. The cosmic rays emitted by these objects can be studied near the earth. This work should improve our understanding of the dynamics of strong gravitational fields and will push general relativity to its limits.

In both the cosmological explosion and the stellar-galactic ones, nuclear, plasma, and even solid-state physics must be employed under physical conditions so extreme that elementary particles could be present with rest masses larger than any now contemplated. Therefore, to understand these phenomena, physicists may have to press other branches of physics to their limits as well.

The discoveries we have mentioned have generated interest among both scientists and laymen. Educated laymen are fascinated by bizarre tales of stars so dense that light waves cannot escape from them, explosions so powerful that a million suns are annihilated in a single stroke, and events taking place so far away in space and time that the universe was only 1 sec old at the time. These things fascinate the scientist, but he is also interested for deeper reasons. In probing regions of curved space we seem to be approaching the problem of creation. Ordinary physics teaches that matter-energy cannot be created or destroyed, but under conditions of the cosmos this rule does not apply in a simple way. In the big-bang theory matter simply appears at a single point in an instant, this event being accompanied by a severe warping of space; in the steady-state theory matter is created by a new kind of force (the "creation field") not yet detected on earth. Notions such as warped space and creation fields might seem remote from earthly concerns, but so at one time the mysterious electromagnetic field conceived by Maxwell must have seemed. This concept later became the basis of all modern communications and of a vast industrial enterprise. It is too early to predict the practical consequences of discovering the precise relationship between general relativity and the real astronomical universe, but historical precedent suggests that they could be substantial.

2.5 PRESENT U.S. RESEARCH EFFORT

The United States is pre-eminent in several branches of astrophysics and relativity. Leadership in this subfield resulted largely from the efforts of the dedicated individuals who raised private and state funds for the construction of the great optical telescopes on the West Coast and of the dedicated scientists who successfully exploited these instruments for galactic and extragalactic research. The optical work on the expanding universe, radio galaxies, quasars, pulsars, and x-ray sources was done largely by these telescopes. These facilities and the cadre of scientists that uses them are a precious national resource that should be carefully conserved and, if possible, strengthened and extended. The two 150-in. optical telescopes now under construction represent encouraging steps in this direction.

Because other countries developed radio astronomy more rapidly than did the United States following World War II, many of the early discoveries in this field occurred overseas. In the early 1960's, as a result of sustained and generous federal support, the United States pulled abreast in radio astronomy. Research on radio galaxies, quasars, and pulsars has been extensive and of high quality. Although the nation now has several fairly large fully steerable dishes for radio astronomy (a 140-ft, a 130-ft, and a 120-ft), as well as larger but not fully steerable dishes (a 1000-ft and a 300-ft), there has been a delay in obtaining support for the large radio instruments recommended several years ago by the Whitford Committee of the National Academy of Sciences* and the Dicke Panel of the National Science Foundation.† Some of these instruments are greatly needed for work in cosmology.

Space astronomy offers unique opportunities for observing high-energy quanta from stellar and galactic explosions. Here the U.S. space program has played a very positive role, with major discoveries in x-, gamma-, and cosmic-ray astrophysics to its credit. Infrared astronomy is largely a U.S. innovation and will be of increasing importance to the study of relativistic

* Panel on Astronomical Facilities, A. E. Whitford, Chairman, *Ground-Based Astronomy: A Ten-Year Program* (National Academy of Sciences–National Research Council, Washington, D.C., 1964).

† Report of the Ad Hoc Advisory Panel for Large Radio Astronomy Facilities, R. H. Dicke, Chairman (National Science Foundation, Washington, D.C., 1967).

objects. Theoretical astrophysics has strong groups at several universities and is at least on a par with the effort in Britain, the Soviet Union, and other countries.

2.6 BENEFITS OF A CONTINUED STRONG PROGRAM

The overall U.S. program in astrophysics and relativity is full of vitality. Should the nation decide to do so, it is in a good position to exploit the field fully and to participate in the exciting discoveries that are certain to take place in the coming years. What are some of the benefits of continuing a strong U.S. program in relativistic astrophysics?

First, because general relativity is a fundamental theory of nature, testing it fully in the solar system and pushing it to its limits in exploding stars and galaxies would prove—or disprove—a theory that forms part of the foundation on which all physical science rests. Moreover, the variety of physical phenomena encountered in cosmic explosions and the extreme conditions under which they occur suggest that surprises and new insights into other fields of physics are likely. For example, cosmic-ray physics and radio and x- and gamma-ray astronomy prove that cosmic explosions can accelerate particles to relativistic energies with surprisingly high efficiency. It is still a mystery how this occurs within our present understanding of plasma physics. If we find out how it occurs, the implications for plasma physics generally, and for controlled thermonuclear fusion in particular, may be substantial.

Second, if a definite cosmological model could be demonstrated to be correct, it would exert a unifying influence on all of science. A valid model would simplify knowledge of the universe, thus making it more accessible. For example, the evolution of the galaxy, the sun, the earth, and life could be placed in their proper context, with a definite sequence of events accounting for the emergence of life and of man.

Third, there would be benefits to education. Not only would philosophy and culture be enriched by a new world view and its impact on considerations of the fate of man, but, on a more direct level, people could follow, through popular articles, a branch of scientific work that has always had broad public appeal. Further, outstanding PhD students trained in this subfield would acquire a broader outlook, which would be relatively well matched to the requirements of applied physics, should they choose the latter field of research.

Fourth, there would be significant contributions to technology. Since objects of cosmological interest are invariably faint, the instruments used to observe them tend to strain to the utmost our technological ability to detect faint sources of radiation and measure them accurately. At the present time there is a highly developed capability in this area, based on strong supporting optical, electronic, and space technology. Maintaining and developing this capability could be a significant advantage in meeting national goals. For example, if the nation should desire at some future time to initiate a project to detect electromagnetic signals from a possible distant extraterrestrial civilization, this ability would be of critical importance. If such signals exist, they undoubtedly will have profound implications for all of human culture.

2.7 RECOMMENDATIONS

In general, substantial benefits will accrue to the vigorous and orderly development of astrophysics and relativity in the United States. The nation has the required scientific and engineering base for such development. Therefore, the Panel believes strongly that full exploitation of the research opportunities in this subfield should continue. We considered a variety of specific steps that would foster this development. When these steps involve the construction of major instruments, the views of the Panel should be regarded as only a part of the rationale for such instruments. Therefore, we have not tried to design a comprehensive program of instrument construction but have confined our efforts to evaluating the instruments that already have been proposed. We have considered their usefulness to astrophysics and relativity, realizing that the decision to recommend the construction of specific facilities should rest with the scientific community and should be based on a broader range of relevant factors. Extended discussion of each recommendation is found in subsequent chapters of the report.

2.8 RECOMMENDATIONS REQUIRING A RELATIVELY SMALL EXPENDITURE

The Panel considered a number of ways in which present facilities and operations can be updated and made more useful for observing objects of interest to astrophysics and relativity. These recommendations to the Physics Survey Committee appear in this section.

Recommendation 1: Scheduling of Large Optical and Radio Telescopes at National Observatories. Cosmological observations often involve faint objects at great distances, so that large amounts of observing time for objects scattered over the sky are required. To plan and to carry out such observations efficiently, investigators need to have a commitment far in advance for the use of the required instruments on a regular basis, even if this means some loss of short-range flexibility for those investigators. *Therefore, we recommend that substantial amounts of observing time on large telescopes at National Observatories (including the 150-in. optical telescopes under construction) should be committed as much as a year in advance to investigators with outstanding long-range programs in cosmology.*

Recommendation 2: Monitoring Variable Radio Sources. The variability of many extragalactic radio sources at millimeter and centimeter wavelengths is important in understanding the physics of these objects. The sources should be monitored on a daily basis. This task can be accomplished by a moderate-size antenna equipped with sensitive receivers and rapid data-processing facilities. *Therefore, we recommend that a moderate-size antenna be equipped for monitoring variable extragalactic radio sources. If an existing antenna is not available, the construction of a new antenna may be required for this purpose.*

Recommendation 3: Very-High-Resolution Studies. Compact radio sources in quasars and radio galaxies appear to be the early stages of a relativistic expansion containing enormous energy; they are very significant for the understanding of the energy-production mechanisms in these objects. Studies with very-long-baseline interferometers show that such sources can be resolved into fine details with antenna systems having a resolving power of the order of 0.001 sec of arc, thus yielding important physical parameters of the source. *Therefore, we recommend that existing radio telescopes be equipped as terminals of a very-long-baseline interferometer for very-high-resolution studies of compact radio sources. The NASA network of tracking stations, containing as it does several long baselines between different stations, and being equipped with low-noise receivers, is well suited for this work and should be made available for it on a part-time basis.*

Recommendation 4: Monitoring Variable Optical Objects. Several types of optical objects of interest to relativistic astrophysics, including x-ray sources, supernovae, quasars, and galactic nuclei, are variable in the optical wavelength range. The time dependence, spectral characteristics, and polarization of these variations can yield significant physical information. To monitor such objects takes a significant fraction of the observing time on an intermediate-size telescope. *We therefore recommend that a number of intermediate-size telescopes be made available for substantial devotion to such monitoring activities, and that they be instrumented with detectors* *and data-handling devices that are adequate for precise and rapid data recording.*

2.9 RECOMMENDATIONS REQUIRING SUBSTANTIAL INVESTMENTS IN GROUND-BASED FACILITIES

Recommendation 5: Construction of Additional Large Optical Telescopes. Even after the completion of the 150-in. telescopes at Kitt Peak National Observatory and Cerro Tololo Inter-American Observatory, the amount of time available for extragalactic observations on large optical telescopes will be less than that which could be used effectively. Such observations are of critical importance to cosmology. *We therefore recommend that, in proposing a balanced program for optical astronomy, the Physics Survey Committee consider the need for additional large optical telescopes to meet the demand for observing time on cosmological problems.*

Recommendation 6: Electronic Optical Imaging. The speed and linearity of electronic digital imaging systems can greatly increase the efficiency of observing faint extragalactic objects with large optical telescopes by permitting rapid and precise corrections for the effects of the night sky. As the investment in large telescopes is substantial, a significant investment in electronic imaging systems is reasonable. *We therefore recommend that each major telescope used for extragalactic observations be equipped as soon as possible with linear digital imaging devices such as high-gain television systems and solid-state arrays.*

Recommendation 7: Large Radio Array. A radio instrument with a beamwidth of the order of seconds of arc could map the structure of nearby strong radio sources, detect additional nearby weak sources, and study distant sources in spite of the problem of confusion with other sources near them in the sky. Such an instrument would make possible both penetrating studies of the physics of radio galaxies and quasars and statistical studies of the number, flux, and angular size of distant sources. These kinds of studies, which are particularly useful if done at several frequencies, are of great importance for relativistic astrophysics and cosmology. Spectroscopic capability will greatly enhance the power of such an array, by permitting observations of the dynamics of galaxies using the 21-cm and possibly other spectral lines. *We therefore recommend that, in proposing a balanced program for radio astronomy, the Physics Survey Committee take into account the need for a large radio array that can synthesize a beam of the order of* *seconds of arc in a reasonable period of time for study of extragalactic radio sources.*

Recommendation 8: Large Millimeter Dish. Extragalactic radio sources at their most active phases radiate powerfully at the millimeter and short centimeter wavelengths. Additional facilities operating at such wavelengths would discover many new active sources and permit studies of their evolution. *We therefore recommend that, in proposing a balanced program for radio astronomy, the Physics Survey Committee take into account the need for a large antenna operating at millimeter wavelengths for observations of active extragalactic radio sources.*

Recommendation 9: Ground-Based Infrared Telescopes. Exploratory observations with moderate-size telescopes on the ground have disclosed that galactic nuclei emit unexpectedly powerfully in the infrared. The cause of this phenomenon is not known, but its occurrence suggests the existence of energy sources—perhaps relativistic—much more powerful than the thermonuclear energy in ordinary stars. It is vital to extend infrared studies to fainter objects and to assess the spectrum and variability with greater precision if we are to understand the physics of the phenomenon. *We therefore recommend that, in proposing a balanced program for infrared astronomy, the Physics Survey Committee take into account the need for a large-aperture telescope at a very dry site equipped and scheduled for infrared observations of extragalactic objects.*

Recommendation 10: Neutrino Astronomy. The attempt to detect solar neutrinos is critically important for all astrophysics. It is particularly so for relativistic astrophysics because of its implications for the theory of stellar evolution, the helium content of the sun, and the possible variation of the gravitational constant. These problems are intimately related to the determination of the age of the galaxy, the problem of the formation of helium in cosmological models, and the choice between rival theories of relativity, all of which are critical for relativistic cosmology. *We therefore recommend that attempts to detect solar neutrinos be supported adequately until decisive results are achieved.*

Recommendation 11: Gravitational Radiation Experiments. Recent experiments suggest that an enormous flux of gravitational waves could be present in space. Confirmation of the detection of such waves would constitute a crucial test of fundamental assumptions underlying the theory of gravitation. A flux of a magnitude even approaching the reported one would have extraordinary implications for astrophysical processes involving relativistic motions as astronomical objects. *We therefore recommend that experiments to detect gravitational waves and study their astronomical sources be fully supported.*

Recommendation 12: Theoretical Studies. Application of the equations of general relativity to astronomically observable objects is important to verify the correctness of the theory and to clarify the basic processes that are occurring. These processes can be inferred from observation only by theoretical reasoning. This activity requires the efforts of mathematicians and physicists of the highest intellectual caliber, together with the judicious use of the most powerful computers available. *We therefore recommend that individuals and groups doing outstanding theoretical work in astrophysics and relativity be adequately supported and that the most powerful computers be made available to them.*

2.10 RECOMMENDATIONS REQUIRING SUBSTANTIAL INVESTMENTS IN SPACE-BASED FACILITIES

Recommendation 13: Diffraction-Limited Optical Space Telescopes. Ultimately the efficiency of observing faint optical objects from the ground is limited almost entirely by atmospheric conditions. In principle this problem can be overcome by a diffraction-limited optical telescope in space. To be competitive with large ground-based telescopes, a space telescope must also be large. A large diffraction-limited space telescope, if it could be built within budgetary limitations, would be of great value to cosmology and relativistic astrophysics because of its ability to study faint, therefore distant, objects, thus permitting the determination of distances of galaxies with sufficient precision to yield the scale and curvature of the universe. *We therefore endorse design studies now under way directed toward flying a large diffraction-limited telescope in space, and we recommend that, as such studies proceed, the effectiveness of the space telescope for extragalactic observations be constantly assessed, as part of the budgetary process, to permit comparison with the effectiveness of contemporary ground-based telescopes.*

Recommendation 14: Infrared and Submillimeter Observations from Aircraft, Balloons, Rockets, and Satellites. If the interpretation of the cosmic microwave background as the effect of a primordial fireball is correct, its intensity should peak near 1 mm and decrease at shorter wavelengths. Confirmation of the predicted variation at and below 1-mm wavelength is an important test of big-bang cosmology. Experiments for this purpose must be performed from high in or above the atmosphere, where its disturbing effects to not mask the background radiation. Preliminary experiments suggest that the background could be more intense than predicted. Similarly, deviations from isotropy could be of great cosmological significance, and observations to find such deviations must be continued.

Recent observations of discrete sources at wavelengths shorter than 1 mm, including infrared wavelengths, indicate that some extragalactic objects radiate most of their power there. It is important to discover the source of such power and to see whether the additive effect of many such sources could possibly account for a substantial portion of the observed background radiation.

We therefore recommend that, as part of an overall program of space observations, a vigorous program of infrared and submillimeter observations of extragalactic background and sources be pursued from aircraft, balloons, rockets, and satellites.

Recommendation 15: High-Energy Astronomy Observatory. X-ray astronomy offers an opportunity to detect relativistic particles at their point of origin by means of their synchrotron and inverse-Compton emission. The origin of enormous quantities of these particles in galactic sources such as the Crab pulsar and extragalactic sources such as radio galaxies and quasars is believed to be connected with violent events. X-ray observations could be decisive in disclosing the nature of these events.

The diffuse background at x- and gamma-ray wavelengths is apparently cosmological in origin. Future observations of its spectrum and angular structure could disclose whether it is truly diffuse, therefore intergalactic in origin, or perhaps the result of the addition of a large number of powerful x-ray sources.

The High-Energy Astronomical Observatory (HEAO) proposed by NASA would permit decisive contributions to the study of galactic and extragalactic discrete x-ray sources and of the diffuse x-ray background by achieving large increases in collecting area, angular resolution, and spectral resolution. The usefulness of the HEAO would be enhanced if a guest investigator program similar to the successful programs operating with other scientific satellites proves feasible in this case as well. In addition, continued rocket and balloon research in high-energy astronomy will be necessary even after the HEAO becomes operational, in order to try out both scientific and technological innovations and to train students. *We therefore endorse the program for the development of a High-Energy Astronomical Observatory, and we recommend that an early start be made on a mission with a large grazing-incidence telescope, capable of high angular and spectral resolu-*

tion. We also recommend that balloon and rocket research in high-energy astronomy be continued at a reasonable level.

Recommendation 16: Gamma-Ray Detectors. The gamma-ray region from 0.5 to 30 MeV, in which extragalactic nuclear gamma-ray lines may be found, is of particular interest to cosmology because of the light it will throw on explosive nucleosynthesis of the heavy elements in distant galaxies, the rate of which depends on the cosmological model. *We therefore endorse efforts to improve detectors in the range from 0.5 to 30 MeV and recommend that the best available instruments in this energy range be incorporated in High-Energy Astronomical Observatory payloads.*

Recommendation 17: Testing General Relativity. The choice between rival theories of gravitation cannot be made conclusively on the basis of present data. This choice is of fundamental physical significance. Moreover, work in relativistic astrophysics depends critically on this choice. *We therefore recommend that experiments using optical, radio, and radar methods to observe the deflection of electromagnetic waves by the sun, the retardation of such waves passing the sun, the precession of the perihelia and apsides of bodies orbiting the sun at various distances, and a possible lengthening of the orbital periods of such bodies be supported and emphasized within a well-balanced program of ground-based and space-based astronomy.* In addition, the use of artificial satellites to detect the inertial drag and geodesic precession in earth orbit should be supported.

3 The Impact of Cosmology on Culture and Science

Historically, only affluent societies have studied cosmology with sufficient intensity to produce definitive advances. Such study—like the rapid accumulation of capital—is one way that strong societies lay a foundation for future greatness.

How cosmology influences society is not well understood; it acts in concert with many other forces in subtle and complex ways. Herbert Butterfield, a historian who tried to evaluate the scientific revolution of the seventeenth and eighteenth centuries, the trademark of which was the cosmology developed by Newton, concluded that, "Since the rise of Christianity, there is no landmark in history that is worthy to be compared with this." The implication is that Newtonian cosmology influenced social forces even more powerfully than did the discovery of America. But it was only at the time of Newton that cosmology linked with the forces that revitalized world civilization. Other examples suggest that this confluence was no accident. There have been three great eras of progress in cosmology—that of classical Greece, that following Newton, and that following Einstein. One could hardly imagine much of our present civilization without the schools of thought that these eras symbolize.

The contribution of Greek cosmology was the idea of making theoretical models to explain what we see—the novel idea of sketching in the mind's eye pictures of things so grand that no human eye could ever actually see them. All men have looked at the heavens with wonder and curiosity, but most found a mythological explanation adequate and returned to mundane occupations. The Greeks were the first to develop scientific theories: The earth is a ball; the moon is illuminated by the sun; each planet moves around the earth in its own orbit. These descriptions were so clear and precise that their inadequacies were provocatively apparent even though they were more profoundly true than any previous explanation of the lights we see in the sky.

In the Newtonian picture of the world, which still concentrated on the solar system, Copernicus placed the sun at the center and Kepler showed how the planetary orbits could be described with mathematical precision. From this basis, Newton was able to show that Kepler's orbits could be derived from physical laws of universal validity, expressed in mathematical form. This achievement inspired the incubation of modern science; it showed the essential role of advanced mathematics and of observations and experiment based on advanced technology and craftsmanship. Furthermore (and especially by the time Newtonian planetary orbits had been confirmed in detail), it inspired and reinforced an unprecedented confidence and ambition to subject all aspects of the world to human understanding and control. According to Butterfield: "When we speak of Western civilization being carried to . . . Japan in recent generations, we do not mean Graeco-Roman philosophy and humanist ideals, we do not mean the Christianising of Japan, we mean the science, the modes of thought, and all that apparatus of civilization that were beginning to change the face of the West in the latter half of the seventeenth century."

Cosmology following Einstein is currently producing a new picture of the universe that is as great an innovation as those of the Greeks or the Newtonians. The new picture is not yet complete, even in outline, but some parts are clear. Problems are studied on a scale that was previously inconceivable. The universe is no longer the solar system surrounded by a scattering of stars, for the stars are seen to be organized into our Milky

Way galaxy, which is one among countless galaxies populating the universe. Furthermore, the galaxies are moving apart in a motion that cannot have been eternal, so cosmology must describe not only the present universe but also its history and evolution. The motion of the galaxies poses the problem of the origin of the universe in an entirely new way that could be directly related to the fundamental structure of matter and space. The problem of the motion of galaxies also forces physical science to take a historical and evolutionary approach to its subject.

Cosmology has a fascination for both laymen and scientists—we are all awed by the mystery of creation. The grandiose scope and fundamental nature of cosmology make it a favorite point of contact and communication between the scientists and the general public. Physicists and astronomers delight to see their own specialties applied to cosmological problems. In a broader sense, cosmology as an observational science now depends on the education of a broader public that is willing, through their elected representatives, to support the allocation of the large sums that are required to advance in this area. Not only is communication with the public a fascinating task for the scientist, it is vital for the progress of his science.

How is it possible to frame cosmology in scientific terms when the universe is unique? This question has particular force when we reflect that in astronomy one can only observe rather than perform controlled experiments as in physics. Usually the fact that various phenomena are observed in great multiplicity compensates in part for this deficiency. Having observed some spectral feature in one star, the astronomer checks many similar stars to see if the feature is typical; having suggested a hypothesis to explain the properties of this feature one can make predictions for stars of very different age, mass, or composition and then observe those stars. Cosmology, on the other hand, deals with the nature and evolution of the universe as a whole, and we know only one universe! It is not surprising that scientists should attempt to approach the mystery of creation by rational hypotheses, but it may be surprising (and gratifying) that such hypotheses can be tested on scientific and not merely aesthetic or philosophical grounds.

The most basic cosmological observations are that the universe is uniform when averaged over distances large compared to the dimensions of clusters of galaxies and that it is expanding. The uniformity is confirmed by the observed isotropy of the cosmic blackbody radiation. The expansion appears to be centered on the observer; however, it is of the type that would seem to all observers, wherever they are located, to be centered on them, so that all parts of the universe are expanding in the same manner. The expansion probably began about 10 billion years ago, and the most direct hypothesis assumes an origin of the universe in a short-lived, high-density, high-temperature phase—the so-called "big bang." In stark contrast to the big-bang model, the steady-state model assumes that continuous creation of new matter replaces the matter lost by expansion in such a way that the universe keeps a constant density throughout all time. Other models take intermediate positions, such as supposing an oscillating universe or a succession of "little bangs."

The scientific testing of various models depends on the fact that information arriving from distant parts of the universe is delayed by the finite speed of light, so that we see galaxies as they were at various times in the past. Thus, in the big-bang model, distant galaxies should appear to be young; while in the steady-state model, galaxies at all distances have the same distribution of ages. The demonstration that the age of a galaxy is correlated with its distance would be decisive evidence against the steady-state model.

Each cosmological model is designed to be compatible with the observed expansion of the universe. Thirty or 40 years ago few additional observations were available, and it was not possible to choose among a variety of models. New observational techniques and refinements of old ones during the last 25 years have converted cosmology into an observational science, that is to say, theoretical models are discarded and new ones invented by comparing their predictions with observations. This comparison, in turn, has stimulated more ambitious and practical-minded theoretical efforts (aided by electronic computers) and further innovations in observing techniques (see Section IV). Specific examples such as the microwave background radiation, which is better explained by the big-bang model than by the steady-state model, or optical data on bright galaxies, which seem to favor an oscillating universe, are found in Section III.

In addition to being a science in its own right, cosmology also impinges on physics in various ways. There are some areas, such as relativity, in which the two subjects overlap. The testing of general relativity, both of the original Einstein formulation and of more recent rival versions, is discussed in Section VI. In a few cases, astrophysics and cosmology have a direct impact on the physics of small-scale phenomena; more often, there is an indirect impact when fundamental physics is forced to deal in a concrete way with natural but unusual phenomena. This subject is discussed in Chapter 4.

4 Past, Present, and Future of Astrophysics and Relativity

4.1 COSMOLOGICAL MODELS

To construct a theoretical model of the universe, we must know what mixture of matter and radiation it contains; we must know the field equations of gravitation; and we must specify the symmetries of the cosmos. Cosmologists are far from agreement on these matters, but the majority view is roughly as follows:

1. The energy content of the universe is now dominated by nonrelativistic matter having neglible pressure. Observations show that galaxies have random velocities of the order of only a few hundred kilometers per second, less than the speed of light by three orders of magnitude. The known populations of relativistic particles, the cosmic rays and photons, have energy densities much less than that of the matter of galaxies.
2. Einstein's equations of gravitation are correct. (The evidence for this view is discussed in Section VI.)
3. All spatial directions and positions are essentially equivalent. The best evidence for the first assumption (isotropy) comes from the observed properties of the cosmic microwave radiation, discussed below. If the universe is isotropic around us, then it is isotropic about every point (unless man is located at a special point—an assumption to be rejected), which implies homogeneity.

These three assumptions lead to a class of cosmological models, first examined in 1922 by Friedmann. In these models, the space–time geometry of the universe is curved, with a scale length that changes with time. As the scale length increases or decreases, galaxies move apart or toward each other, like dots painted on the surface of a balloon that is being inflated or deflated. Light from distant galaxies therefore undergoes a Doppler shift toward the red, if the universe is expanding, or toward the blue, if it is contracting, and these wavelength shifts increase with distance. The Friedmann models thus provide an explanation of the systematic red shifts of distant galaxies, first observed during the 1920's.

The Friedmann models are distinguished by just two numerical parameters, which can be taken as the present rate of expansion, known as the Hubble constant, H_0, and the present rate at which this expansion is slowing down, known as the deceleration parameter, q_0. In all these models, the universe began with an infinitely high density at a time in the past less than the Hubble time $1/H_0$, and it has been expanding ever since. If q_0 has a value greater than one half, then the scale of the universe is now expanding more slowly than the two thirds power of the time and will eventually stop expanding and contract to a state of infinite density. The deceleration parameter is proportional to the present mass density of the universe and is less than, equal to, or greater than one half, depending on whether the mass density is less than, equal to, or greater than a critical value determined by the Hubble constant. Also, if q_0 is less than or equal to one half, the universe is infinite, with negative or zero spatial curvature, although, if q_0 is greater than one half, the universe is spatially finite, with positive spatial curvature. Thus, as long as we adhere to the Friedmann models, the chief task of observational cosmology is to determine the critical parameters H_0 and q_0.

Before describing the progress made in this task by observational cosmology, we must indicate the various minority views that would modify the Friedmann models in one way or another. The least radical modifications are those that alter the material contents of the universe or the gravitational field equations but leave intact the assumptions of homogeneity and isotropy. Walker and Robertson showed in 1935 that all such models have essentially the same space–time structure as the Friedmann models, except that the cosmic-scale factor can have quite a different time dependence. Included in this class of models are the following:

1. Models with a Cosmological Constant When Einstein first turned his attention to cosmology in 1919, he assumed that the universe is static. However, no static solutions of his field equations could be found, so he modified them by adding a new term, which involves a new constant, the so-called cosmological constant. The original motivation for a cosmological constant was removed by the discovery of the expansion of the universe, but its existence remains a logical possibility. Models with a cosmological constant can have quite different histories from Friedmann models and need not begin with a state of infinite density.

2. The Steady-State Model Bondi, Gold, and Hoyle have gone beyond spatial isotropy and homogeneity and have assumed that the universe, although expanding, always looks the same (temporal homogeneity). To fill the widening gaps between galaxies, new matter must be continuously created, but the steady-state model is not specific as to how and where this occurs. In this model, the cosmic scale factor grows exponentially. The Hubble constant is really a constant of nature (unlike the factors of the Friedmann models, which decrease with time), and the deceleration parameter q_0 takes the value -1. The universe is spatially flat, but the space–time continuum is curved like the surface of a sphere of radius $1/H_0$ in a five-dimensional flat space.

3. Models with Massless Particles It is possible that the energy density of the universe is dominated not by galaxies that we see but by particles of zero rest mass, such as neutrinos or gravitons, which interact too weakly with matter to have been detected. Another possibility is that the field equations involve massless scalar fields, as in the theory of Brans and Dicke. The resulting models differ from the Friedmann models only in detail and, in particular, share thė feature of an initial state of infinite density.

A more radical step is to give up the assumption that the universe is homogeneous and isotropic. Its space–time structure now no longer has the simple form derived by Robertson and Walter; everything is much more complicated. An example is the hierarchical model of Charlier. Stars are grouped into galaxies, galaxies form clusters, and there is even evidence that clusters of galaxies are grouped into superclusters. Charlier suggested that this hierarchy continues indefinitely. In such models there is not even any meaningful way of averaging out its properties to define isotropy or homogeneity. The weight of present opinion is that hierarchy terminates with clusters of clusters of galaxies and that the universe is isotropic and homogeneous on any larger scale.

With this wide range of possible cosmologies, why do astronomers usually analyze their data in terms of the Friedmann models? The reason is not that these models are surely right, but that we need a very restrictive theoretical framework to draw any quantitative conclusions about the geometry and history of the universe from the limited data now provided by observational cosmology. As the resources of astronomy improve, we may well find that the Friedmann models no longer fit the data. At that point, it will be appropriate to adopt a less restrictive framework, either by incorporating a cosmological constant into the field equations, by postulating continuous creation, by allowing the presence of a scalar field or a high proportion of massless particles, or by giving up the principles of isotropy or homogeneity.

The Friedmann models present a challenge to observational astronomy: to determine the Hubble constant, H_0, and the deceleration parameter, q_0. Once these two parameters are known, the Friedmann models will tell us how old the universe is, whether it will continue to expand forever or begin to contract again, and whether it is spatially infinite or finite. However, the effort to determine H_0 and q_0 could reveal that the Friedmann models are wrong; then we would be able to develop a more accurate cosmology.

The problem of determining H_0 and q_0 has called into play virtually every weapon in the armamentarium of astronomy. In rough order of importance, these include the following:

1. Red Shifts and Luminosities In a uniformly expanding universe, the velocity away from us of a moderately distant galaxy is its distance times the Hubble constant, H_0. H_0 is often expressed in years^{-1}. The velocity of a galaxy can be accurately measured by observing the red shift of its spectral lines, and the distance can be determined by observing the apparent luminosity of the galaxy, which decreases with distance according to the inverse-square law. We also need to know the absolute luminosity of the galaxy, or the energy that it radiates. This is accomplished by an elaborate chain of distance determinations, which is continually being checked and refined. In a recent determination of H_0, the chain consisted of five links. The value of H_0 is subject to change as the data are refined.

First, the distance to a nearby cluster of stars, the Hyades, is determined from motions of its stars to be 130 light-years. Many stars of the Hyades cluster belong to the main sequence, the class of stars that have not yet exhausted the hydrogen fuel at their centers. The absolute luminosity of such stars is known to be uniquely correlated with their surface temperature (spectral type). Knowing the distance to the Hyades, we can deduce for each spectral type the absolute luminosities of its main-sequence stars from their apparent luminosities.

Second, the distances to other stellar clusters and associations within our galaxy are determined by comparing the apparent luminosities of their main-sequence stars with the absolute luminosities known from the Hyades for the same spectral type. Six of these clusters and associations contain a total of nine giant variable stars, called Cepheid variables. The absolute luminosity of such stars is uniquely correlated with their period of variation and spectral type. Knowing the distance to the nine Cepheids in open clusters and associations, we can deduce their absolute luminosities from their apparent luminosities.

Third, the distances to the nearby galaxies within our local group are determined by comparing the apparent luminosities of their Cepheids with the absolute luminosity for the same period and spectral type. It is found that the distance to M31, the great spiral galaxy in the constellation Andromeda, is two million light-years. Knowing the distance to M31, we can determine the absolute luminosity of its globular clusters—great agglomerates that contain hundreds of thousands of individual stars.

Fourth, the distance to the nearest rich cluster of galaxies, in the constellation Virgo, is determined by comparing the apparent luminosity of the brightest globular cluster in the Virgo galaxy M87 with absolute luminosity of the brightest globular cluster in the Andromeda galaxy M31. If one assumes that these brightest globular clusters have the same absolute luminosities, the distance to M87, and hence to the Virgo cluster, is 50 million light-years. Knowing the distance to the Virgo cluster, we can determine the absolute luminosities of its constituent galaxies from their apparent luminosities.

Fifth, the study of a number of rich clusters of galaxies led Hubble to conclude that all their brightest members have about the same absolute magnitude. If one assumes that the brightest galaxy of any cluster has the same absolute luminosity as the brightest galaxy M87 of the Virgo cluster, then measurement of the apparent luminosity of the brightest galaxy in a cluster provides a measure of the distance to the cluster.

Using this particular chain of distance determinations, $1/H_0$ was found to be between 10 billion and 16 billion years. However, any change at any link in the chain of cosmic distance determinations would require a corresponding change in the Hubble constant, so the true range of uncertainty in H_0 is probably even wider; recent work suggests that $1/H_0$ is somewhat larger.

To determine q_0, it is necessary to push the measurement of red shifts and luminosities to such great distances that the graph of galaxy velocity versus distance is no longer a straight line. The curvature of this graph, together with the Friedmann models, then determines q_0. The most distant galaxy known is the radio source 3C 295, for which the fractional red shift of wavelengths is 46 percent. The available data for other galaxies out to 3C 295 indicate that q_0 is between one half and three halves. However, the measurement of q_0 has its own difficulties, including sensitivity to evolutionary and selection effects. The steady-state model behaves like a Friedmann model with $q_0 = -1$ and does not allow evolution, so it is in apparent conflict with the measurements of red shifts and luminosities.

2. *Age Measurements* According to the Friedmann models, the age of the universe should be less than the Hubble time, $1/H_0$. On the other hand, the age of the galaxy can be estimated from the known present relative abundances of the isotopes of uranium. If one assumes that uranium was created during a short period (lasting not more than a few hundred million years), with ^{235}U initially about 65 percent more abundant than ^{238}U (as indicated by considerations of nucleosynthesis), then for the ^{235}U to have decayed to its present low abundance, the uranium must have been formed about 7 billion years ago. If the uranium was created gradually, our galaxy must be older than that, and estimates based on the abundance of other radioactive species give ages ranging up to 20 billion years. The comparison of the observed distribution of globular cluster stars in luminosity and spectral type with the results of theoretical calculations of stellar evolution gives an age between 9 and 15 billion years. The upper range of these age determinations is larger than the lower range of determinations of the Hubble time, but no clear discrepancy has yet emerged.

3. Mass Density Measurements If the Hubble constant is (13 billion years)$^{-1}$ and the deceleration parameter is greater than one half, then the mass density of the universe in a Friedmann model must be greater than 9×10^{-30} g/cm^3. The masses of a few nearby galaxies are known from studies of their rotation velocities or the dynamics of pairs of neighboring galaxies. From the mass-to-luminosity ratios of these galaxies, and the observed density of luminosity throughout the universe, the mass density of the universe can be estimated to be about 3×10^{-31} g/cm^3. The study of clusters of galaxies gives a slightly larger result, but it still appears that the mass density of luminous matter is too small by an order of magnitude to be consistent with a q_0 as large as one half. It is possible that nonluminous intergalactic matter makes up the missing mass. If such matter were cool, it would produce absorption effects in the light from distant quasars, and such effects have not been seen. If it were hot, the absorption effects would not be expected, but then one might see emission effects, for example at x-ray wavelengths. There is inconclusive evidence that such effects exist. Therefore, the estimates of mass density that can be supported by direct observation suggest that q_0 is considerably less than the critical value of one half. This would imply an open, indefinitely expanding universe.

4. Angular Diameters, Number Counts The measurement of angular diameters of galaxies as a function of their red shifts could in principle serve to determine H_0 and q_0, but galaxies vary too widely in actual diameter to make this method practicable. The measurement of the numbers of galaxies up to a given red shift or down to a given apparent luminosity (or radio source strength) might also be used to determine q_0, but the radio source counts do not agree with any Friedmann model, unless we allow for evolutionary effects. Since the reason for such effects is unknown, this method is useless at present as a measure of q_0. Again, the steady-state model, which does not permit evolution, appears to be in conflict with the source counts. The influence of selection effects could be greater than we have assumed, in which case this method is useless.

In summary, the confrontation of the Friedmann models with different pieces of astronomical evidence leads to different pictures of the universe, of which the following are examples:

1. The universe is finite and the deceleration parameter is of order unity, as indicated by the measurements of red shifts and luminosities. Therefore, the bulk of the matter of the universe must be in some form that has not yet been discovered, possibly intergalactic ionized hydrogen with a temperature of order 10^6 K or greater. The universe is relatively

young, with an age close to the lower limits allowed by the dating of radioactive elements and globular clusters.

2. The universe is infinite, with a very small deceleration parameter, as indicated by the observed density of luminous matter. The age of the universe is close to the Hubble time, in rough agreement with other estimates.

It is also entirely possible that neither of these alternatives represents the true Friedmann model, or even that the Friedmann models are altogether wrong.

4.2 NUCLEOSYNTHESIS IN THE BIG-BANG AND STELLAR EXPLOSIONS

A challenging problem is to account for the observed abundances of the elements, which seem to be very similar in the solar system, in distant stars of the galaxy, and even in other galaxies. Gamow and his collaborators developed the original big-bang model because they thought that the elements could be formed under the intense heat and high density of the early phases occurring in Friedmann models. Although light elements can form in this way, all theoretical attempts to produce nuclei heavier than lithium have failed. In a steady-state model, in which there is no hot, dense phase, one is forced to consider nucleosynthesis in stars. Thus, for the heavy elements, both types of model require stellar nucleosynthesis.

For the light nuclei (hydrogen, deuterium, helium, and lithium), synthesis in the big bang is a possible explanation under certain simple assumptions. If Einstein's general relativity is the correct theory, hydrogen and helium would be produced in a ratio of about ten to one by number, a value approximately equal to that actually observed in most objects, and deuterium and ^{3}He would be roughly 10^{-4} and 10^{-5} of hydrogen, respectively. However, if the scalar–tensor gravity theory is correct, it is more likely that little helium would be produced. Astronomy still has not settled this most important question: Did the helium exist when galaxies began to form or has it been synthesized by nuclear reactions in stars within the galaxies? The steady-state model demands the latter, ordinary big-bang models demand the former, and scalar–tensor big-bang models could yield either. The most natural place to look for an answer, the composition of the oldest stars, offers some indications either way; thus the matter is not yet settled.

Evidence regarding the big bang could also come from detection of deuterium and ^{3}He by radio astronomy. Observation of the hyperfine transition in $^{3+}$He in H II regions would allow its abundance to be determined. Similar searches for the hyperfine transitions of interstellar deuterium have been unsuccessful so far, but the corresponding upper limit is not decisive.

A series of papers by Cameron, Fowler, Hoyle, and the Burbidges developed the idea that nucleosynthesis of elements heavier than helium occurs in the interiors of stars. By surveying the systematic properties of nuclei, these papers proposed a set of thermal environments of various initial compositions that would produce, by thermonuclear reactions, the prominent features of the natural abundance distribution. Recently, it has become clear that the environments yielding the best reproduction of the abundances are not found in the slow evolution of stars but, rather, in dramatic last-second explosions—supernovae and "little bangs," which extensively alter the composition of stellar matter. It now seems possible that all the elements and their isotopes with atomic number greater than or equal to $Z = 6$ (carbon) have been synthesized during the explosions of massive stars (roughly 20 to 40 times the mass of the sun). The evidence that the elements were synthesized at the moment of explosion comes from comparing the nuclear abundances theoretically produced in the two cases with the actual nuclear abundances observed in nature. The key feature is that in an explosion nuclear fuels burn at temperatures considerably higher than those at which the same fuels burn in a static star. The final abundances are different because of the higher ignition temperature in the explosive case. Inasmuch as the big-bang model appears to be incapable of providing for the synthesis of elements heavier than lithium, it is gratifying to find that, at least theoretically, stellar explosions can produce them.

It has long been realized that the atomic nuclei cannot always have existed. The fact that radioactive ^{235}U, with a half-life of 0.7 billion years, still constitutes nearly one percent of the much longer-lived ^{238}U assures us that much of the uranium on earth was produced in a few-billion-year period immediately preceding the formation of the solar system; if the uranium had been produced much earlier, the ^{235}U would have decayed by now. It is likely that the uranium was produced in a nearby supernova explosion. Presumably nucleosynthesis is still occurring, and, if it is, we can expect to detect the fresh radioactivity by techniques of gamma-ray astronomy. The best gamma rays to search for are those emitted when ^{56}Ni decays to ^{56}Fe after its ejection from exploding stars. The evidence is now quite good that abundant ^{56}Fe was synthesized in that manner, and the observed average density of iron in the universe suggests that the associated density of nuclear gamma rays also should be observable. We are faced with the exciting possibility that gamma-ray astronomy at photon energies between 0.5 and 4 MeV may provide hard experimental facts concerning the history of nucleosynthesis in the universe. A single observation of the ^{56}Ni gamma-ray spectrum will throw light on these subjects: (a) the occurrence of explosive nucleosynthesis in the universe, (b) the dependence of its rate on cosmic time, and (c) the validity of the steady-state universe.

The most direct way that the study of stellar evolution and nucleosynthesis provide information regarding the correctness of cosmological models is in the timing of the creation of the elements. In Friedmann models the age of the universe is less than the reciprocal of Hubble's constant, or 10 billion to 16 billion years. If we live in a Friedmann big bang, that age should considerably exceed the age of the globular clusters, about 9 billion to 15 billion years as estimated by customary calculations of stellar evolution. Because the globular clusters are metal poor, their age must considerably exceed the mean age of the elements as revealed by the radioactive probes. The most farseeing of these probes, ^{187}Re, seems at present to be also about 15 billion years old, although different ages have been advanced by different scientists. The uncertainty in each of these three ages is still sufficiently great that there is not necessarily a contradiction, but the suggestion of one has been lurking for years, and we early await several proposed measurements that will help to clarify these ages.

4.3 EXPLODING GALAXIES

Until about 25 years ago, practically all of our information about the universe was obtained from optical studies. With increasingly sophisticated techniques, optical astronomers were investigating the properties of our own galaxy and the expanding universe outside. Basic understanding of the energy sources and structures of the stars and the dynamics of the galaxy were obtained. Although 20 years before, the universe had been shown to be expanding, few data relevant to cosmology were available. The investigation of the electromagnetic spectrum in the postwar period—first by the techniques of radio astronomy and more recently in x-ray and gamma-ray wavelengths and in the infrared—and the increasing realization that classical cosmic-ray physics involving charged particles is closely interwoven with these parts of the electromagnetic spectrum, led to a new era of discovery. Major discoveries, largely unpredicted, have occurred throughout the last two decades.

Radio astronomers first showed that there are large numbers of sources in the universe that are radiating vast energies in the radio spectrum between about 10^7 and 10^{11} Hz. The form of the spectra, the polarization, and other evidence led quite early to the conclusion that these sources are emitting by the electron synchrotron process (fast electrons in magnetic fields). The theory of synchrotron radiation implies that they must be generating very large energies (up to 10^{61} ergs equivalent to 10^7 solar masses in some cases) of relativistic particles. Some sources are in our own galaxy; others are extragalactic objects. The sources in our galaxy are remnants of exploding stars—supernovae—the most famous one being the Crab nebula. The early optical studies of this object showed that it was an old supernova remnant, and the discovery that it was a gigantic particle accelerator led to a revival of the idea that supernovae are the primary sources of cosmic rays. The extragalactic sources were even more remarkable. The vast energies present must be released in gigantic explosions, usually located in the nucleus of the object. Theoreticians are still trying to account for this vast energy release. The amounts of energy are so large that it appears that only the release of gravitational energy through collapse of a large mass in a very small volume, or the creation of matter, is able to account for these phenomena. The gravitational fields present are likely to be strong, requiring general relativity for their interpretation.

More recent discoveries have added to this picture. It has been shown that the nucleus of our own galaxy is a powerful source of nonthermal radio emission and also of infrared emission. The major source of activity is confined to a region with a size of only one or two light-years (compared with the size of the galaxy, which is 100,000 light-years). Similar situations have been found in galaxies of practically all types. Recent investigations of infrared emission in some galaxies show that it is much greater than all the radiation emitted by the stars and is generated in very small nuclear regions. Radio astronomers, using interferometers with very long baselines (approaching the diameter of the earth), have measured components with these very small sizes (as small as 0.001 sec of arc) directly in distant galaxies. Moreover, there is now ample evidence that the nonthermal emission in optical, infrared, and radio wavelengths, which is generated in the nuclei of galaxies, often shows variations in times of the order of years, months, and possibly days, indirectly confirming that the energy is generated within volumes less than a light-year across. The gravitational radiation detected by Weber, discussed in Section 5.7, may be a manifestation of even greater activity in the nucleus of our galaxy, provided that the claimed detection is confirmed.

Thus, studies on radio, optical, and infrared wavelengths have led us to realize that stars and starlight are not the only energetic constituents of the visible universe. Violent activity in the nuclei of galaxies is probably commonplace and is not due to thermonuclear processes.

So far we have discussed only the previously unknown characteristics of galaxies revealed by investigations outside the optical wavelength band. Even more significant are the entirely new kinds of objects that have been found. The most important of these are the quasi-stellar objects (quasars).

and the pulsars (see Section 4.8). Prior to 1960, it was found that the optical counterparts of extragalactic radio sources were galaxies, objects having discernible angular diameters (unlike stars). In 1960, some radio sources were identified with pointlike objects indistinguishable from stars on direct photographs. At first it was thought that the sources might be like true stars, hence the name quasar (which is short for *qua*si-stell*ar*). A major breakthrough came in 1963 when it was found that the spectrum of the brightest quasar, 3C 273, has a large red shift ($\Delta\lambda/\lambda = 0.16$). In the ten years that have elapsed, extensive work on quasars has been conducted, but their nature still remains a mystery. Their red shifts extend from about 0.1 to about 2.9 (distances from 10^9 to 10^{10} light-years, if Hubble's law applies). Their optical, radio, and infrared spectra are in many ways similar to those of the nuclei of galaxies in which violent activity is taking place, and the simplest apparent interpretation of them is that they are indeed excessively luminous galaxies either in an early stage of evolution or in a very late stage. But there is so far no direct evidence that they are actually galaxies. If they are at cosmological distances, there are many severe problems associated with the idea that they are galactic nuclei with scaled-up energies. Some scientists therefore have proposed that they are not at great distances, so that their red shifts are not of cosmological origin. Attempts to establish conclusively that most quasars are at cosmological distances have failed, as have attempts to argue that they are comparatively close by, that is, at distances of less than 10^8 light-years. This situation is currently at an impasse, the majority of scientists tacitly assuming that they are at the distances suggested by their red shifts, so that they may be used for cosmological investigations; while the minority feel that, without an independent distance determination and an understanding of the physics of the objects, it is premature to apply them to test cosmological models in ways that might be completely irrelevant.

In any case, quasars are of extreme importance for relativistic astrophysics. If they are at cosmological distances, they are the only discrete objects that we can now use directly to investigate the universe as it was billions of years ago. If they are closer to us, then their existence means that comparatively stable configurations must exist that can give rise to large intrinsic red shifts, possibly casting doubt on the reliability of the red shift of ordinary galaxies as a distance indicator. In either case, we are very far from understanding the way in which the particles are accelerated, although large masses inside small volumes, hence strong gravitational fields, must be involved.

Although quasars were discovered by the radio astronomers, it is now clear that radio emission is sometimes a comparatively minor form of activity in such objects. Optical objects with similar properties but less radio emission are a major constituent of the universe. It is probable that the time over which they are active is short compared with 10^{10} years. This assumption implies that the total number of objects that are their progenitors or descendents could be comparable to the total number of luminous galaxies in the universe.

4.4 RADIO GALAXIES AND COSMOLOGY

The discoveries of radio galaxies and quasars posed enormous theoretical problems because of the huge energies involved. Moreover, because they are so bright, they offer the possibility of extending cosmological studies to much larger red shifts, where the effects of curvature also are large.

Surveys of the sky at radio frequencies show numerous discrete sources. The brightest of these are concentrated near the galactic plane and are clearly in our Milky Way. The others are nearly uniformly distributed about the sky and are extragalactic. Modern radio telescopes have sufficient sensitivity and resolution to isolate about one million such sources. Of the few hundred sources with radio positions good enough to permit optical study, about one third are identified with galaxies and one third with quasars. For the remaining third, no optical identification has been possible, sometimes because of heavy obscuration by dust, because the optical luminosity is relatively low, or because the galaxy is very distant.

Since the early 1950's, scientists have realized that many of the discrete sources are sufficiently strong to be detected by radio means, although they are far beyond the limits of even the largest optical telescopes. As a striking example, the most distant known galaxy is identified with a radio source, 3C 295, which, as a radio source, is 10^4 times stronger than the weakest source that we can detect. Presumably many of the faint radio sources recorded correspond to optical galaxies too faint to see. Although the potential applications to the cosmological problem are obvious, no definitive answers have been found so far. The main difficulty is the inability to determine from radio measurements alone the distance and red shift of a radio source. Clearly the discovery of a "standard candle" for radio astronomy, analogous to the "first ranked cluster member" used in optical investigations, would be of very great value, as would a feature in the radio spectrum that could be used to find the red shift. But so far it has been necessary to depend on an optical magnitude and red shift. Since detailed studies are necessarily limited to those objects that can be studied optically, the potential power of the radio telescope in investigating the very distant parts of the universe has not been fully realized.

For this reason most of the radio-astronomy effort in cosmology has depended on statistical studies of relative numbers of radio sources of various intensities and angular sizes, based on the reasonable assumption that in general the most distant objects will appear fainter and smaller. Unfortunately, the spread in intrinsic luminosity and linear dimensions of the extragalactic sources is very large, and this has greatly limited the certainty of the results. The simplest interpretation of the data indicates a large excess of faint, presumably distant, radio sources, over what would be expected in a universe uniformly filled with radio sources. It is generally assumed that this excess indicates a very much greater density of strong radio sources at large distances. Because of the travel time of light, this great density at large distances corresponds to larger densities at early epochs. Such a theory is inconsistent with the simplest form of steady-state cosmology, which requires the universe to be unchanging with time. Some astronomers, however, challenge the cosmological significance of the source counts and choose to interpret the data as merely indicating a deficiency of radio sources in the local region, rather than an excess at large distances. Moreover, the experimental result is not without criticism. Even if one accepts an excess density at large distances, a detailed understanding of the situation is not possible without first knowing the distribution of intrinsic luminosities, which requires knowledge of the distance to a large number of sources.

Some attempt has been made to divide the radio counts into categories according to type of associated optical object, surface brightness, or radio spectrum to determine which, if any, particular class of source is responsible for the observed excess. So far the results have been inconclusive, primarily because of the small number of sources involved and the difficulty of interpreting the effect of observational selection introduced when one tries to classify or identify radio sources. A major need is for more systematic data.

There is today a growing realization that, before any definitive cosmological conclusions can be drawn from radio studies, it will be necessary to understand better the nature of the radio sources themselves. Even aside from their use as a tool for cosmology, the extragalactic radio sources stand as a major problem for theoretical astrophysics. The release of gravitational energy of very massive bodies has appeared to be one of the most promising energy sources proposed and has stimulated extensive theoretical research on massive bodies and gravitational collapse. Other suggestions, including matter–antimatter annihilation, quarks, or a "creation field" have equally profound implications for fundamental physics.

4.5 X-RAY AND GAMMA-RAY SOURCES

The potentially observable range of the electromagnetic spectrum extends over 25 decades of frequency from megahertz radio waves to high-energy gamma rays. More than half of these decades lie above the ultraviolet region of the spectrum and constitute the province of x- and gamma-ray astronomy. Though solar x rays were first observed in 1948, it was only during the 1960's that the region of extrasolar x and gamma rays was scouted in experiments with rockets, balloons, satellites, and ground-based instruments. Beginning with the discovery of the first x-ray source in 1962, a variety of galactic and extragalactic sources has been uncovered. Extreme physical conditions exist in these sources that were never before encountered in terrestrial physics or astrophysics. They constitute, in effect, distant laboratories where one can observe previously inaccessible regimes of plasma and nuclear physics. Moreover, the observations of extragalactic x and gamma rays provide new data about the conditions in intergalactic space that are of fundamental importance to cosmology.

The first x-ray source, Sco X-1, is now known to be a variable stellar object emitting over 99.9 percent of its radiation in the form of x rays whose spectral distribution is characteristic of a plasma with a temperature of about 50 million degrees. Its average x-ray luminosity is over 1000 times the total luminosity of the sun. It flickers irregularly and occasionally flares by factors of 3 or more within minutes. It has been suggested that Sco X-1 is a close binary system undergoing violent mass exchange, but no evidence of the expected orbital periodicity has been found yet. Imbedded in the Crab nebula is a rapid pulsar (see Section 4.8), discovered by radio observations and subsequently observed in the optical and x-ray regions. Over 99 percent of its pulsing radiation is in the x-ray region at energies above 1 keV.

Dwarfing these galactic sources in total x-ray luminosity is the galaxy M87. This giant star system is distinguished by its conspicuous nonthermal radio emission and by a peculiar chain of beadlike features (the "jet") that emit polarized continuum radiation and extend about 4000 light-years from the nucleus. Rocket experiments have shown that M87 emits x rays at a power level that exceeds its radio power and is a considerable fraction of the optical luminosity of all its stars. A challenge for future observations is to determine whether the x rays emanate from the "beads" or are emitted by the compact radio source in the nucleus. Their polarization is also of interest. As in the case of the Crab pulsar, we may be witnessing in these beads the effects of rotating magnetized bodies, but on an enormously larger scale involving millions of solar masses. Other extragalactic objects, including the Seyfert galaxy NGC 4151, the radio galaxy Cygnus A, and the quasar 3C 273, have been identified as x-ray sources. It will be interesting to find out whether the x rays in these sources are produced by the relativistic particles that are known to be present from the radio emission observed.

In the earliest rocket observations of galactic x-ray sources, the presence of a diffuse component of energetic cosmic photons was inferred from the existence of an apparently uniform background intensity. Many subsequent observations from rockets and balloons have established that this background is grossly isotropic and therefore probably of extragalactic origin. Its spectrum shows several distinctive features that appear to have fundamental cosmological significance. Below 1 keV, where the measurements are complicated by the effects of absorption in our galaxy, there is evidence for the existence of a particularly high intensity of soft x rays, which has been interpreted as the free–free emission of a universal intergalactic gas, with a kinetic temperature of the order of one million degrees. An abrupt change in the slope of the spectrum occurs near 40 keV. One theory attributes this finding to a corresponding change in the spectrum of electrons in intergalactic space, which are thought to produce the observed x rays by inverse Compton collisions with starlight and the cosmic microwave background (see Section 4.7).

Recent observations suggest that the x-ray background is not perfectly isotropic but is associated with clusters of galaxies. If this interpretation is correct, an important new tool for exploring distant cosmic matter will have been gained.

Measurements in progress may determine if the low-energy gamma-ray spectrum near 1 MeV shows any evidence of the decay of ^{56}Ni, produced in supernova explosions during an earlier era of rapid star formation in the evolution of the universe. A resolution of the basic question of whether the observed background is truly diffuse or is the combined effect of many distant discrete x-ray emitting galaxies awaits the development of instruments with improved angular resolution and sensitivity. In any case, the properties of the extragalactic high-energy photon flux are a basic part of the observational material with which cosmology must now deal.

High-energy cosmic gamma rays were observed for the first time in a satellite experiment launched in 1967. Energetic secondary gamma rays must be produced when cosmic-ray protons collide with interstellar matter; therefore, they can serve as a probe of the distribution of high-energy interactions in the galaxy and universe. An instrument aboard the OSO-3 satellite detected gamma rays with energies of about 100 MeV arriving, as expected, from a band of directions coincident with the galactic disk in which cosmic rays and interstellar matter are concentrated. A strong peak of gamma-ray intensity from the direction of the galactic center is, however, an unexpected feature of the observations that awaits a satisfactory explanation. In addition to the galactic gamma rays, the observations give evidence for an isotropic intensity of gamma rays, which appears to be a part of the diffuse extragalactic photon flux.

Attempts have been made to detect cosmic gamma rays in the ultra-high-energy range above 10^{11} eV by observing from the ground the extensive air showers they produce in the atmosphere. So far only upper limits on the intensity of various potential discrete sources have been obtained. Nevertheless, since the spectrum of primary cosmic rays extends to at least 10^{19} eV, one can assume that the spectrum of secondary cosmic gamma rays extends to comparable energies, though the intensity is too low to be detected with present techniques. In this high-energy region the spectrum of extragalactic gamma rays may show the effects of the opacity of intergalactic space, which is expected because of photon-photon collisions between the gamma rays and low-energy photons in starlight and in the cosmic microwave background. These opacity effects should appear near 10^{13} and 10^{16} eV for starlight and microwave background, respectively.

4.6 COSMIC RAYS

The study of cosmic rays within our galaxy is important to cosmology because it provides detailed information about the fast particles that are the sources of energetic photons. A comparison between the visible properties of our galaxy and other galaxies then permits inferences about the general characteristics of extragalactic radio and x-ray sources.

Cosmic rays with energies between 10^{10} and 10^{14} eV are studied by direct observations near the earth. Energies much below 10^{10} eV are strongly affected by the solar wind. Information in the low-energy range may be obtained by observing the secondary effects of such particles in interstellar space, such as heating of the interstellar gas and production of x rays by "knock-on" electrons. Above 10^{14} eV, one may study effects produced in the earth's atmosphere, such as the extensive showers of secondary particles produced when a high-energy cosmic ray collides with a nucleus high in the atmosphere.

The directional distribution of the charged cosmic rays is highly isotropic. At first this finding seems inconsistent with the assumption of a galactic origin for them. In contrast, the anisotropic distribution of electromagnetic radiation clearly reflects its predominant origin in the galactic disk. However, energetic charged particles accelerated in one way or another in the galactic disk can be trapped within the disk by the galactic magnetic field. Their trajectories are so twisted by magnetic force that any detectable trace of the anisotropic distribution of their sources is destroyed. Thus, up to at least 10^{16} eV, which appears to be about the energy limit for effective galactic trapping, it is plausible to look within our galaxy for the sources of cosmic rays. The discovery of pulsars has thrown new light on the origin of cosmic rays within our galaxy. The pulsars are apparently efficient transformers for converting gravitational energy into high-energy cosmic rays. The accelerated particles are retained within the galaxy by the galactic magnetic field for times of the order of 10^7 years before they collide with interstellar matter or leak into intergalactic space. The rate at which energy is apparently being converted by pulsars in this way could be sufficient to supply all the cosmic rays in the galaxy.

In view of the probable origin of most, if not all, of the galactic cosmic rays in pulsars, we can look on the charge composition of the primary nuclei as essentially a chemical analysis of the material in the immediate vicinity of the pulsar, if not the pulsar itself. The composition is, of course, modified by the passage of the nuclei through the interstellar matter. In recent years, the techniques for measuring this composition have advanced rapidly, so that detailed analysis of the abundance ratios of nuclei differing in Z by only one, and even isotopic composition studies among the low-Z elements, has become feasible. This is particularly exciting because certain of the observable isotopes are radioactive. By studying the isotopes whose half-lives against decay are in the right range, it is hoped that one can determine the length of time the cosmic rays are trapped in the galaxy. When used together with the known number of cosmic rays trapped in the galaxy, this lifetime yields a production rate required to account for their presence—and this is an important constraint on models of cosmic-ray sources. Ultrahigh-Z nuclei have been observed, and strong evidence for transuranic nuclei has been reported. The study of fossil tracks of cosmic-ray nuclei in crystalline lunar rocks has provided information on the intensities of galactic cosmic-ray nuclei for several billion years.

The energy spectrum of primary electrons—both positive and negative—has been studied intensively. Electrons in the energy range from 1 to 10 GeV, which are responsible for much of the galactic nonthermal radio emission, are now known to be primarily negative. Since the secondary electrons from cosmic-ray interactions with interstellar matter would contain roughly equal numbers of positrons and negatrons, one can conclude that most of the 1- to 10-GeV electrons are directly accelerated. This conclusion is evidently consistent with a pulsar origin. On the other hand, the proportion of positrons increases toward lower energies, indicating that they are of secondary origin with an increasing contribution from collisions of protons and nuclei.

Primary nuclei with energies greater than 10^{19} eV have been detected by observing the extensive showers of particles that they produce when they strike the atmosphere. Since the magnetic rigidity of such nuclei is so great that they cannot be contained by the magnetic field of our galaxy, they are almost certainly of extragalactic origin. However, photons of the cosmic microwave background interact with protons above a threshold energy near 10^{20} eV to produce pi mesons. Since the mean free path for such interactions is much less than the Hubble radius, the volume of space within which the observed extragalactic component of cosmic rays can originate must shrink at high energies. Therefore, one can anticipate a cutoff in the spectrum somewhere above 10^{20} eV. Experiments now in progress can check this prediction and thereby provide a significant new test of our ideas concerning the origins of high-energy cosmic rays and the universality of the blackbody radiation. Some evidence has been found to show that the primaries above 10^{19} eV differ from those below 10^{14} eV in being either nearly pure protons or nearly pure heavy nuclei instead of a mixture. Since the highest energy primary cosmic-ray nuclei are, in effect, intergalactic messengers, it would be extremely interesting to know, in addition, whether any of them are actually antinuclei produced perhaps in antigalaxies. Unfortunately, no feasible method exists at the present time by which one can distinguish between nuclei and antinuclei at ultrahigh energies.

4.7 COSMIC MICROWAVE BACKGROUND

The first theoretical treatment of a radiation-filled universe was given in 1931, and in 1948 big-bang cosmology was used in an attempt to account for the production of heavy elements by nuclear reactions in the hot material. Although this proved to be unsuccessful as a way of producing appreciable amounts of elements heavier than helium, Gamow was able to estimate a present temperature of 5 K for the blackbody radiation, a fossil remnant of the big bang.

Based on measurements with a 1.25-cm radiometer, an upper limit of 20 K and isotropy better than ±1 K were obtained for this radiation in 1945. More interesting is a determination of 2.3 K, which could have been obtained in 1941 from a measurement of the population of the first excited rotational state of cyanogen in interstellar space. The first positive indication of the radiation, a radiometer measurement giving an excess temperature of 3.5 ± 1 K at 7 cm, was made at Bell Laboratories in 1965. The interpretation of these observations as cosmic blackbody radiation was stimulated by a project then under way at Princeton to construct a radiometer at 3.2 cm to look for such radiation. Subsequently, a temperature of 3.0 K was obtained at 3.2 cm. In 1968, much better measurements gave a temperature of 2.7 K at wavelengths of 3.2, 1.6, and 0.86 cm. The shortest wavelength superheterodyne radiometer measurement is at 3.3 mm. By analyzing spectroscopically the excitation of cyanogen, a temperature of 2.8 ± 0.2 K has been obtained at a 2.6-mm wavelength. Particularly important for cosmology are the measurements of isotropy. These show that at wavelengths of the order of 1 cm the universe appears to be remarkably isotropic, and, furthermore, that the earth's velocity relative to this isotropic universe is at most 300 km/sec. Further studies of the isotropy of the background radiation may yield information on the magnitude of irregularities in the early universe.

An anomalously high temperature was reported at wavelengths below 1 mm using a bolometer in a rocket and tentatively confirmed by one balloon measurement. More recent balloon measurements give intensities closer to those expected on the basis of a 2.7 K blackbody spectrum. The significance of this anomaly, if real, is not yet clear. If the excess radiation at these short wavelengths proves to be correct, and is observed to be isotropic with a continuous spectrum, it will be necessary to assign a cosmic origin; but in the absence of these observations, the radiation could have its origin in the galaxy. If the origin should be cosmic, interesting questions about the theory of the fireball and the origin of the radiation at longer wavelengths would be raised.

A host of interesting theoretical developments has grown from the discovery of the cosmic microwave background. Some of these, such as the question of helium formation in the fireball of the early big-bang universe and the formation of galaxies through the growth of instabilities, have their roots in work by Gamow and collaborators in the late 1940's. The expected fractional yield of helium computed using general relativity, about 27 percent by mass, is remarkably insensitive to neutron half-life and present matter density, although it does depend on the net lepton number. On the other hand, the scalar–tensor theory can require very small helium formation dependent on the unknown strength of the scalar field.

The present state of our observational knowledge of the amount of helium formation that actually occurred in the fireball is not in good order. Generally, nuclear reactions in stars are expected to increase the helium content found in the interstellar medium and in young stars. Consequently, one must look at old stars to estimate the abundance of helium in primordial material. But most old population II stars are too cool to show helium lines. The evolved population II "blue horizontal branch" stars are hot enough and show very weak helium lines corresponding to a low surface abundance. The significance of this result is somewhat unclear because of the existence of some weak-lined young blue stars with weak helium lines. According to general relativity, high helium abundance seems to be required if the age of population II stars is to be less than the age of the universe, but under the scalar–tensor theory low helium is required to give a reasonable age after including the evolutionary correction caused by early rapid burning. Among other possible sources of information concerning primordial helium are the quasars, which apparently have weak helium lines, and the low yield of solar neutrinos, which seems to require somewhat low solar helium; the direction of evolution of the stars in the horizontal branch of the Hertzsprung-Russell diagram of globular clusters; the helium-dependent mass and luminosity of stars on the horizontal branch in comparison with the calculated luminosity and calculated mass–period relation of RR Lyrae variables; and the luminosity of population II stars with dynamically determined masses suggest nearly normal helium. None of these means of assessing the primordial helium abundance has given a completely convincing answer, and the problem continues to be a matter of great importance.

Much remains to be done. Observations at wavelengths below 2 mm are exceedingly difficult, but this part of the spectrum is very important. A great deal of energy could be contained in it, and the high-frequency tail of the spectrum is capable of providing important tests of the fireball hypothesis and the theory of the thermalization of the radiation. Many more balloon and rocket observations will be required to provide a clear picture of this spectral region. Another set of interesting questions involves the interstellar molecular lines. Are optical pumping or collision-induced transitions affecting the rotational populations, thus yielding anomalous temperatures?

4.8 PULSARS

Pulsating radio sources, or pulsars, were discovered by accident in a scintillation study of low-frequency radio sources. Over 60 of these extraordinary objects have been discovered. Characteristically they emit sharp pulses of radiation about once per second. The timing of these pulses is so precise that one speaks of millisecond accuracy over time intervals of a month or more. A particularly interesting pulsar is that embedded in the Crab nebula. It emits pulses simultaneously over the entire electromagnetic spectrum, from radio frequencies to x rays, at a rate of 30 times per second. It is generally thought that the Crab pulsar is the collapsed, superdense, stellar remnant of the same supernova explosion that was observed in A.D. 1054 at the present position of the Crab nebula. This remnant is probably a neutron star about 10 km in diameter, which is spinning at the pulse frequency and is gradually slowing down as its rotational energy is used in accelerating particles to cosmic-ray energies. Its electromagnetic radiation, probably synchrotron in origin, is emitted in a beam that sweeps across the earth like the beam of a distant lighthouse.

In the initial state of the collapse of the stellar core, gravitational potential energy is converted into rotational energy of the spinning neutron star. In the process, the star's magnetic field is compressed and amplified to 10^{12} G or more. Such a strong field, when rotating with the neutron star, generates enormous electromotive forces capable of accelerating electrons and ions to 10^{16} eV or more. The electrons emit the radiation we see, and the ions are a prime source of cosmic rays.

Pulsars are the latest in a series of unexpected and spectacular discoveries of modern astrophysics. The strong emission of radio bursts at highly regular intervals, of the order of a second, was unexpected because of the high intensity of this radiation, but the possibility of stars rotating with millisecond periods had been suggested on theoretical grounds. In fact, the possibility of neutron stars, with diameters of the order of 10 km and densities of the order of nuclear density, had been suggested by theoretical physicists as soon as neutrons had been discovered. Rotation periods of as little as a millisecond have been predicted, whereas all other known stellar objects (including white dwarfs) have rotation and vibration periods in excess of a second. The rotation of three white dwarfs has been found to be slow. A number of pulsars have now been observed with periods much shorter than a second, and for many the period has also been found to increase a measurable amount in a year or so. As bizarre as the idea of a whole star at nuclear densities may sound, the central energy source of a pulsar must be a highly condensed object, with strong gravitational fields. Rotational energy of neutron stars is considered the most likely—and the most conservative—hypothesis at the moment.

At least some pulsars—such as the one in the Crab nebula—are known to be remnants of supernova explosions, which is gratifying to the theorists, since they had speculated that these explosions are preceded and triggered by the dynamic collapse of an already dense highly evolved star. This collapse may lead to a highly condensed central remnant with strong gravitational fields (possibly relativistic); pulsars are presumably the outward manifestations of such remnants. The theoretical study of pulsars invokes almost every major branch of modern theoretical physics. The structure of a neutron star depends on relativistic equations of state for matter consisting mainly of neutrons and hyperons; quantum-mechanical phase transitions apply to the charged components—free protons versus aggregates of complex nuclei. General relativity is of some importance for the understanding of stable neutron stars and of overwhelming importance for fully collapsed objects. Even for neutron stars of low mass, the gravitational binding energy is much larger than the nuclear binding energy of even the most stable nuclei in the iron region. Even solid-state physics is relevant, in spite of temperatures far in excess of a million degrees, since the complex nuclei can form a "Coulomb lattice" at the extreme densities inside a neutron star. There are now such detailed and complex observations on changes in pulsar periods that theorists talk in earnest not only of solids but of starquakes, of superconducting currents, and even of volcanoes.

The rich structure of the interior of a rotating neutron star is matched by the complexities theorists invoke for their exteriors to explain the radio emission (not to mention strong gravitational radiation that is predicted for a young, distorted, and rapidly rotating pulsar). Rival theories abound for dealing with the finer details of the emission mechanism, but it is fairly clear that particle acceleration, relativistic plasmas, and coherent emission phenomena all play a role. From hyperons in their deep interior to cosmic rays in their magnetosphere, pulsars provide a testing ground for modern physics (or at least for the imagination of modern physicists). Pulsar theories also have stimulated new kinds of theoretical models for quasars. Rapidly rotating, massive, dense disks now seem attractive, possibly with a "vanished Schwarzschild singularity" in the center and, in any case, with strong gravitational fields.

4.9 PROSPECTS FOR FURTHER ADVANCE

Recent work embodies two somewhat contradictory tendencies. On the one hand, some of the predictions of relativistic models have been verified, such as the existence of what appears to be a blackbody background. On the other hand, some observations have cast doubt on these predictions. For example, the possible existence of noncosmological red shifts in quasars would confuse the straightforward interpretation of the Hubble expansion. Such a dichotomy is not unexpected in this field, because we are dealing with complex phenomena, which, although energetic in absolute terms, are usually very distant and therefore faint and hard to interpret. Often observations are undertaken to test one or another prediction of a model, and, although partial confirmation is frequently the result, just as often the increase in sensitivity required to make the test discloses new phenomena previously unsuspected. For example, in an attempt to extend radiogalaxy observations to large red shifts, radio astronomers discovered quasars. Strong radio galaxies tend to have nearly uniform optical luminosities and therefore have been useful as "standard candles" in extending the Hubble relation to large red shifts. Quasars are much more luminous, thus potentially useful to larger red shifts, but their luminosity varies from source to source. This variation makes them useless for extending the Hubble relation until the physics of the source is better understood.

Future work in the subfield will probably have a similar character. It will be stimulated by the desire to solve well-posed problems. Conceivably, the answers to these problems will fit a simple pattern, and a definite model for the evolution of the universe will triumph. More likely, no simple pattern will emerge. Unsuspected phenomena probably will be discovered, and revisions of the theoretical picture will be necessary. In either case, however, we will be coming closer to answering the question, "How did the universe originate?"

The United States is in a good position to make a major contribution to this subfield in the next few years. The nation has large optical telescopes, advanced radio telescopes, and space telescopes, as well as excellent theoretical groups working in both relativity and astrophysics. Other nations, notably Great Britain, the Soviet Union, Germany, The Netherlands, and Australia, are also in a good position to contribute, primarily as a result of efforts in radio astronomy and theoretical astrophysics, and several have indicated their intention to do so. Our country should continue its earlier efforts and share in the discoveries that will be made.

As we look to the future, we see three major resources that can be brought to bear:

1. The established groups of astronomers, with their present instruments, and the groups of theorists working in relativity and astrophysics;
2. A reservoir of physicists who, attracted by the opportunities of this field, will begin to apply to it their own specialties in other subfields of physics (as this reservoir is about ten times the number already in the field, a relatively small percentage shift can make substantial changes in the rate of progress);
3. Frontier instruments now under discussion that should be built (we refer here particularly to large ground-based and space telescopes that can extend the range at which we can detect galaxies and relativistic objects).

It is probable that each of these resources can and will be brought to bear. In Chapters 5 and 7 we discuss their application and potential.

5 Research Methods in Astrophysics and Relativity

5.1 OPTICAL METHODS

Large optical telescopes have played a key role in studies of the red shift–magnitude relation of the brightest members of clusters of galaxies. They are necessary to detect galaxies of large red shift, which are distant and therefore faint. The Mount Wilson 100-in. telescope yielded a red shift of 13 percent for the Bootes cluster as early as 1936. No further progress in extending the red shift–magnitude relation was made until the completion of the Palomar 200-in. telescope. Soon after the prime-focus spectrograph was finished, a red shift of 20 percent was determined for the Hydra cluster. Larger red shifts—up to 46 percent for 3C 295—have since been determined for radio galaxies with strong emission lines in their spectra.

In studies of stellar evolution, large optical telescopes have made decisive contributions. Stellar evolution is best studied in star clusters whose members were formed at essentially the same time. Evolution will affect the brightest stars first, but to calibrate the observed effects, it is necessary to study the fainter stars that still show no evolutionary effects. These fainter stars allow a determination of the cluster's distance and an estimate of its chemical composition. In some crucial globular clusters in our galaxy, observation of stars of the required faintness with a 200-in. telescope requires a substantial part of a clear, moonless "good-seeing" night for each star. Photons arrive very slowly.

Spectroscopic studies yield the chemical composition of stars, which reflects the nucleosynthetic history of the interstellar gas from which they were formed. These studies require a large light-gathering power and a high spectroscopic dispersion, which can be obtained only in large stable spectrographs. The coudé spectrographs of the large telescopes have produced essentially all the information that we have concerning the chemical composition of stars.

The aperture of an optical telescope determines its light-gathering power and its angular resolving power. However, the effective resolving power of large ground-based telescopes is, at good locations, limited to about 1 sec of arc, on the average, by turbulence in the earth's atmosphere. Only high in the atmosphere or above it is a gain in resolving power obtained for larger apertures.

The main attribute of a large optical telescope is its light-gathering power. Specifically, an efficient photoelectric spectrometer at the 200-in. telescope allows detection of one photon per second; such a photon intensity corresponds to a star whose brightness is 18th magnitude (corresponding to a galaxy with a red shift of 0.1) over a spectral bandwidth of 5 to 10 Å. The reliable detection of emission or absorption lines in the spectrum depends directly on the total number of photons detected. Since most observational work in cosmology is on faint objects, it is probably realistic to say that the rate of observational progress is determined by the total light-gathering power of the larger telescopes. On this basis, the completion of the two 150-in. telescopes, now under construction at Kitt Peak and Cerro Tololo, will increase light-gathering power available to U.S. astronomers by around 60 percent. The power would be more than doubled if another 200-in. telescope were built in addition.

Cosmological studies frequently require extended observations of many faint objects—a slow procedure. The triple requirements of a clear sky, no moonlight, and "good seeing" contribute to the slowness of progress in these observational programs. Only one sixth of the total night time is suitable for observational work of critical cosmological importance.

Some observatories assign observing time to local and outside scientists on an annual basis, allowing them to plan and to carry out long-term programs. It has not been lack of scientific planning and foresight that has slowed progress in observational cosmology, but rather the lack of large telescopes. If the two new 150-in. telescopes are to contribute significantly to observational cosmology, a sizable fraction of their time should be committed a year ahead, in large blocks, to a limited number of astronomers on submission of outstanding long-term programs.

Beyond the obvious hazards such as clouds and moonlight, the main problems besetting the astronomer observing faint objects are seeing and night-sky emission. The night-sky emission consists of molecular bands, simulating an irregular continuum at low spectroscopic dispersion, and a few strong forbidden atomic lines. Near cities there is scattered continuum and mercury-line emission. The contamination of a star's spectrum by that of the night sky depends on the size of the "seeing" disk, the apparent angular diameter of the star due to atmospheric turbulence. The seeing disk is usually about 1 to 2 sec of arc in diameter at selected observing sites. The night-sky light within the seeing disk is as bright as a star of apparent magnitude 21, making it difficult to detect faint spectral features in stars much fainter than 19th magnitude (such as the sun, were it a faint member of a globular cluster) or in galaxies where surface brightness is added to the light of the night sky.

Since the night sky represents the ultimate limitation in earthbound observations, the performance of an instrument used in observing faint objects is characterized, first, by the time required to detect the night sky and, second, by the feasibility of subtracting the effect of the night sky from the observed spectrum of the faint object. The time to detect the night sky is essentially measured by the optical efficiency of the spectrograph and the quantum efficiency of the detector. The detector efficiency has improved considerably in recent years through the replacement of the photographic plate (efficiency of about 1 percent) by the photoelectric effect (efficiency of about 20 percent) used in image tubes, television tubes, and photomultipliers. In image tubes, the output is usually recorded on photographic emulsions, with their high two-dimensional resolution ability but nonlinear response. The spectrum of object and night sky appear side by side, but the complicated nonlinear response of the plate makes the correction for night-sky contamination unreliable. The output of a photomultiplier is usually in pulse counts corresponding to photoelectrons, hence the detection is linear; there is no imaging, and thus there is no space resolution. The effect of the night sky can be taken into account by subtracting the counts from a neighboring sky area of a size equal to that of the area centered on the star. Multichannel photoelectric spectrometers that employ a row of detectors allow moderate resolution, but in one dimension only.

The development of systems such as integrating low-noise television cameras, which combine the advantages of high-resolution one- or two-dimensional imaging with linear detection response, is feasible now and, though costly, deserves high priority. The speed of the system is high if coupled to image intensifiers. The linearity ensures proper subtraction of the night sky. The system would increase the efficiency of observing very faint objects by a considerable factor. The ideal instrument would detect each photon that strikes each resolution element. It is unlikely that further large gains are possible for ground-based telescopes, since the ultimate efficiency depends on the quantum efficiency of photoelectric surfaces and on the seeing.

The next large step requires a diffraction-limited large space telescope, for which the contamination of the stellar image by foreground light will be very much less than that for ground-based telescopes. The uses of such a telescope for cosmology have been considered in a recent report of the Space Science Board.* According to this report, a properly figured 120-in. telescope in earth orbit should yield images less than 0.1 sec of arc in diameter. The amount of night-sky light within this disk, which contaminates the stellar image, is reduced both because of the smaller disk

* Space Science Board, Division of Physical Sciences, National Research Council, *Scientific Uses of the Large Space Telescope* (National Academy of Sciences, Washington, D.C., 1969).

and because of the absence, in earth orbit, of atmospheric emission. When these effects are combined, a limiting magnitude of 29 appears to be attainable; however, to achieve this in practice would require improvements in imaging devices at the focus. Such a limiting magnitude is 100 times fainter than can be observed from the ground. Therefore, galaxies or quasars like those that we have studied could be observed at ten times the distance, if the effect of space curvature is neglected. Such capability should permit determination of distances to galaxies sufficient to yield the scale and curvature of the universe.

Most of the efficient auxiliary instruments and sophisticated data-handling systems are, and should be, developed in association with the larger telescopes. These systems improve the efficiency of smaller telescopes as well, allowing observational programs such as the extended monitoring of the beat of the Crab pulsar. For some programs a small telescope is called for, since the image scale at the large telescope is excessive; for example, observations of the energy distribution of nearby giant ellipticals, important for the interpretation of red shift–magnitude diagrams, were conducted with a 4-in. telescope and a spectrum scanner at Palomar.

In conclusion, we see the need for several steps to increase the available capability for observing faint objects with large telescopes. The following four recommendations to the Physics Survey Committee outline these needed steps:

Recommendation 1: Scheduling of Large Optical Telescopes at National Observatories. Cosmological observations often involve faint objects at great distances, so that large amounts of observing time for objects scattered over the sky are required. To plan and to carry out such observations efficiently, investigators need to have a commitment far in advance for the use of the required instruments on a regular basis, even if this means some loss of short-range flexibility for those investigators. Therefore, *we recommend that substantial amounts of observing time on large telescopes at National Observatories (including the 150-in. telescopes under construction) should be committed as much as a year in advance to investigators with outstanding long-range programs in cosmology.*

Recommendation 5: Construction of Additional Large Telescopes. Even after the completion of the 150-in. telescopes at Kitt Peak National Observatory and Cerro Tololo Inter-American Observatory, the amount of time available for extragalactic observations on large optical telescopes still will be less than that which could be used effectively. Such observations are of critical importance to cosmology. *We therefore recommend* *that, in proposing a balanced program for optical astronomy, the Physics Survey Committee consider the need for additional large optical telescopes to meet the demand for observing time on cosmological problems.*

Recommendation 6: Electronic Imaging. The speed and linearity with which one can approach the limits imposed by the night sky can be greatly increased by electronic digital imaging systems. Because of the large investment in telescopes used for extragalactic work and the few nights available per scientist, increasing their efficiency by the use of such systems deserves high priority. *We therefore recommend that each major telescope used for extragalactic observations be equipped as soon as possible with linear digital imaging devices such as high-gain television systems and solid-state arrays.*

Recommendation 13: Diffraction-Limited Space Telescopes. Ultimately the efficiency of observing faint objects from the ground is limited almost entirely by atmospheric conditions. In principle this problem can be overcome by a diffraction-limited telescope in space. To be competitive with large ground-based telescopes, a space telescope must also be large. A large diffraction-limited space telescope, if it could be built within budgetary limitations, would be of great value to cosmology and relativistic astrophysics, as it would permit the determination of distances of galaxies with a precision sufficient to yield the scale and curvature of the universe. *We therefore endorse design studies now under way directed toward flying a large diffraction-limited telescope in space, and we recommend that, as such studies proceed, the effectiveness of the space telescope for extragalactic observations be constantly assessed, as part of the planning and budgetary process, to permit comparison with the effectiveness of contemporary ground-based telescopes.*

5.2 RADIO METHODS

Radio telescopes are limited by sensitivity and angular resolution. There are many problems in which the simplicity of a single antenna is useful, including the investigation of total flux density from known sources as a function of wavelength, polarization, and time. There the primary consideration is sensitivity, and the largest possible aperture is desirable. If, however, one desires information about the spatial distribution within the source, as in the case of radio galaxies in which clouds of relativistic particles are ejected, single antennas often do not provide sufficient resolution. This problem also occurs when one tries to resolve individual faint sources from one another, as in observing distant radio galaxies for cosmological purposes. Then it is necessary to use interferometers or arrays to provide large baselines and high resolution, although not necessarily a large collecting area. By making observations with different baselines it is possible to build up an image of the source.

For many applications, such as the study of very faint extragalactic radio sources, the sensitivity is limited not by receiver noise but by confusion from unresolved sources in the antenna beam. In these cases the required resolution can be obtained only by using very large arrays. As an example, we note that at 20-cm wavelength the newly completed array in The Netherlands can observe sources one to two orders of magnitude fainter than can a 100-m paraboloid (operating at the same wavelengths), which has about the same physical collecting area and cost.

Present equipment in the United States includes a large variety of single dishes, interferometers, and arrays. Most of this equipment had already been constructed or was in the process of construction at the time of the Whitford report in 1964,* and, although recommendations were made in the report in support of a program of advanced new instruments, none of these instruments has been completed. We find that many of the proposals made there for instruments to be applied to cosmology are still generally valid.

Future investigation of the cosmological problem by radio observations will require the sensitivity and resolution of large arrays to measure the size and intensity of a sample of the more distant radio sources. Arrays will allow the detection of now undetected galaxies, quasars, and relatively nearby clusters. This will be invaluable in understanding the relation between galaxies and quasars, discovering why or when a galaxy becomes a radio source, and determining the statistical distribution of radio luminosities for analysis of radio source counts. A large array with spectroscopic capability can be used to study the dynamics of normal and unusual galaxies using the 21-cm and other spectral lines.

Arrays capable of a resolution of 1 sec of arc might permit counts of source number versus angular size as a test of cosmological models. This test is analogous to, but more powerful than, the relation between source number and flux density. Such resolution will permit study of the structure of radio sources, as a function of wavelength and polarization, over a wider range of flux density than now possible. Such investigations may uncover a relation between absolute radio luminosity and some observable

* Panel on Astronomical Facilities, A. E. Whitford, chairman, *Ground-Based Astronomy: A Ten-Year Program* (National Academy of Sciences–National Research Council, Washington, D.C., 1964).

radio property, which would permit estimates of distance from radio data alone. Extended radio sources need to be studied with high angular resolution and sensitivity, for a better understanding of the nature and evolution of strong radio sources.

The fundamental problem of energy sources may be illuminated by study of compact young sources. Temporal variations in intensity and size will determine significant properties, during and immediately following the release of energy and acceleration of relativistic particles. At early epochs these sources radiate most strongly at millimeter and centimeter wavelengths. Two types of instrument are needed: a large antenna to discover and study the details of such sources and antennas capable of monitoring time variations in the strong sources. For the latter only moderate-size antennas are needed, if provided with state-of-the-art receiving and data-processing techniques.

Investigation of the angular structure of very young and rapidly varying sources requires interferometer baselines of thousands of kilometers—very-long-baseline interferometers (VLBI). Existing radio telescopes can be instrumented, at little cost, to become part of an array with transcontinental and intercontinental baselines with sufficient resolution to study the structure and high-velocity variations of radio sources. VLBI patrols have already begun with a growing fraction of time at radio-astronomy observatories devoted to this type of work. Ultimately, special antenna facilities may be desirable to provide different baseline orientations. The worldwide NASA network of tracking stations is ideally suited for VLBI work with several long baselines between widely spaced stations with excellent low-noise receivers.

The discussion of optical methods referred to the special character of cosmological observations, requiring large blocks of time scheduled considerably in advance and possibly for many years of program. This is also true in radio astronomy, because of the faintness and variability of the sources studied.

In summary, we recognize the need for additional instruments that can map the structure of nearby extragalactic sources with high resolution, observe very distant and faint sources in spite of the confusion problem, monitor variable sources at short wavelengths, discover additional sources at millimeter wavelengths, and study small sources with very high angular resolution. Instrumentation to meet these objectives and requirements for its use are described in the following recommendations to the Physics Survey Committee:

Recommendation 7: Large Radio Array. An instrument with a beam width of the order of seconds of arc can map the structure of nearby strong sources, detect additional nearby weak sources, and study distant faint sources in spite of the confusion problem. Such an instrument will make possible both penetrating studies of the physics of radio galaxies and quasars and statistical studies of the number, flux, and angular size of distant sources. Both are of fundamental importance to relativistic astrophysics and cosmology. Spectroscopic capability will add considerably to the power of a large array, by permitting observations of the dynamics of galaxy spectral lines. *We therefore recommend that, in proposing a balanced program for radio astronomy, the Physics Survey Committee take into account the need for a large array that can synthesize a beam of order of seconds of arc in a reasonable period of time, for study of extragalactic radio sources.*

Recommendation 8: Large Millimeter-Wave Dish. Recent investigations show that extragalactic radio sources at their most explosive phase radiate most powerfully at the millimeter and short centimeter wavelengths. Additional facilities operating at such wavelengths would discover many new active sources and permit studies of their evolution. *We therefore recommend that, in proposing a balanced program for radio astronomy, the Physics Survey Committee take into account the need for a large antenna operating at millimeter wavelengths for observations of active extragalactic radio sources.*

Recommendation 2: Monitoring Variable Sources. There is a need to monitor variable extragalactic radio sources at millimeter and centimeter wavelengths on a daily basis. This task can be accomplished with a moderate-size antenna equipped with sensitive receivers and rapid data-processing facilities. *We recommend that a moderate-size antenna be equipped for monitoring variable extragalactic radio sources. If an existing antenna is not available, the construction of a new antenna may be required for this purpose.*

Recommendation 3: Very-High-Resolution Studies. Compact radio sources in quasars and radio galaxies appear to be the early states of a relativistic expansion containing enormous energy; therefore, they are very significant for understanding the energy-production mechanisms in these objects. Studies with very-long-baseline interferometers show that such sources can be resolved into fine and variable details with antenna systems having a resolving power of the order of 0.001 sec of arc, thus yielding important physical parameters of the source. *Therefore, we recommend that existing radio telescopes be equipped as terminals of a very-long-baseline interferometer for very-high-resolution studies of compact radio sources.*

The NASA network of tracking stations, containing as it does several long baselines between different stations and being equipped with low-noise receivers, is well suited for this work and should be made available for it on a part-time basis.

Recommendation 9: Scheduling of Large Radio Telescopes at National Observatories. As in the case of optical telescopes, referred to in Recommendation 1, cosmological investigations require the commitment of substantial blocks of time as much as a year in advance. We therefore expand the scope of Recommendation 1 to include radio telescopes.

5.3 INFRARED ASTRONOMY

Infrared astronomy is still in its infancy, at least in comparison with investigations at optical and radio wavelengths. It was anticipated that many objects in the universe must be comparatively cool, so that the thermal radiation they emit might peak in the infrared. Using modest instruments and comparatively simple detectors, astronomers have already obtained important results in planetary and stellar astronomy by detecting cool objects that radiate powerfully in the infrared. Of particular importance are objects that could be stars in the process of formation and stars that are rapidly evolving. The infrared radiation apparently comes from vast clouds of dust near the object, which are being heated by the central star.

That extragalactic objects would be exceedingly powerful infrared emitters was not anticipated. For example, from observations so far, it appears that the nuclei of some galaxies emit infrared energy between 1 and 20 μm at power levels 50 or 100 times greater than all the stars in that galaxy. Infrared astronomy has already shown that the nuclei of galaxies—apparently objects in which relativistic effects occur—radiate at power levels so high that they could not continue through the life of a galaxy unless extensive mass flows into the nucleus or matter is created there. These results, of great importance to relativistic astrophysics, have been obtained by a few small groups working largely from the ground. Observations from aircraft and balloons, using a 12-in. telescope, show that the nucleus of our own galaxy emits strongly at wavelengths beyond 50 μm.

We have already discussed whether the microwave background is blackbody radiation generated in a big-bang fireball. Observations to settle this question are needed between about 1 mm and 100 μm. At present, some of the direct observations made above or near the top of the atmosphere, using rockets and a balloon, conflict with the blackbody interpretation.

They may conflict with data from interstellar molecules, bathed in the background radiation field. If correct, a major argument for an evolving universe would be eliminated, but the preliminary observations made from above the atmosphere could be wrong. More data are needed to draw any firm conclusions. Critical observations include those of the spectrum at wavelengths below 1 mm and attempts to find deviations from isotropy.

The radiation of so much power in the infrared by certain discrete extragalactic objects is a great puzzle. The mechanism has not yet been identified, but the data point to an extremely large energy source. We must progress far beyond the present fragmentary observations with 60-in. telescopes and with the 120-in. and 200-in. telescopes on Mt. Hamilton and Palomar Mountain. The latter telescopes are located at sites optimized for optical work but not for infrared work.

For progress in infrared observations of extragalactic objects, astronomers need continued access to a large telescope at a very dry site selected for its suitability for infrared work. Such a telescope would be used in all the conventional infrared windows and could also make observations at wavelengths near 1 mm, where windows to be exploited also exist yet. As infrared and millimeter detectors improve, we anticipate that such an instrument will penetrate ever further into space, permitting study of increasing numbers of extragalactic objects. This study, including spectra, variability, and polarization, should disclose the nature and source of the extraordinarily powerful infrared emission. It should be noted that such an instrument would be far less expensive than an optical telescope of comparable size, because of the relaxed tolerances associated with the longer wavelength.

Many parts of the spectrum between 1 mm and 1 μm cannot be observed properly unless the observation is done above, or at least high in, the atmosphere. Most of the power in some objects is in the wavelength range between 100 μm and 1 mm. Some observations at these wavelengths can be made with a large telescope from a dry site. However, the appropriate atmospheric windows are rather restricted. Thus, infrared astronomy must also be pursued from aircraft, balloons, rockets, and satellites. The anticipated gain from more flights involving a 12-in. telescope at 50,000 ft in an airplane, a 36-in. telescope at the same altitude, a 36-in. telescope at 100,000 ft on a balloon, or any satellite telescope with good pointing accuracy is great. No quantitative predictions can be made, because so little has yet been done, but it is obvious that this is an important direction in which to move.

A space program devoted to infrared and submillimeter wavelengths is also necessary to measure the spectrum of the microwave background at wavelengths below 1 mm. The techniques are quite different from those used to measure sources, for, if the background is isotropic as expected, absolute flux measurements (rather than measurements of the amount by which a source exceeds the background) are required. In either case, sensitive detectors must be flown high in or above the atmosphere.

The small but rapidly growing number of infrared astronomers in the United States need more technical support, more money for instrumental development, and worldwide surveys to find the best high-altitude sites on which to construct new infrared telescopes. They also need adequate logistic support to enable them to operate observatories at such sites. Interferometry, at infrared wavelengths, will allow measurement of extragalactic objects of exceedingly small size. We do not know yet the extent to which this technique will be developed, but clearly two infrared telescopes will be required.

Infrared astronomy could hold the key to several baffling problems in astrophysics and relativity. At present it is an exploratory effort involving a relatively small group of scientists; however, it should be expanded, and more advanced instruments should be made available. The following recommendations to the Physics Survey Committee offer specific suggestions:

Recommendation 9: Ground-Based Infrared Telescopes. Exploratory observations with moderate-size telescopes on the ground have disclosed that galactic nuclei are unexpectedly powerful infrared emitters. They are radiating more energy than is available from thermonuclear sources; therefore, they probably involve relativistic effects. Larger-aperture telescopes are needed for studies of the physics of these unusual objects. *We therefore recommend that, in proposing a balanced program for infrared astronomy, the Physics Survey Committee take into account the need for a large-aperture telescope at a very dry site equipped and scheduled for infrared observations of extragalactic objects.*

Recommendation 14: Infrared Observations from Aircraft, Balloons, Rockets, and Satellites. If the interpretation of the cosmic microwave background as the effect of a primordial fireball is correct, its intensity should peak near 1 mm and decrease at shorter wavelengths. Confirmation of the predicted variation at and below 1-mm wavelength is an important test of big-bang cosmology. Experiments for this purpose must be performed from high in or above the atmosphere, where its disturbing effects do not mask the background radiation. Recent observations of discrete sources at wavelengths shorter than 1 mm, including infrared wavelengths, indicate that some extragalactic objects radiate most of their power there. It is important to discover the source of such power and to see whether the additive effect of many such sources could possibly account for a substantial portion of the observed background radiation.

We therefore recommend that, as part of an overall program of space observations, a vigorous program of infrared and submillimeter observations of extragalactic background and sources be pursued from aircraft, balloons, rockets, and satellites.

5.4 X-RAY AND GAMMA-RAY ASTRONOMY

In the past decade, rapid developments in the instruments and methods for x-ray and gamma-ray astronomy have occurred. We now stand on the threshold of a decade in which satellite experiments in high-energy astronomy can reap a harvest of discovery and new understanding in astrophysics and cosmology.

During 1970 and 1971, all-sky x-ray survey experiments with 1° angular resolution were launched on NASA satellites. Based on balloon and rocket experiments through 1967, these experiments have detected hundreds of x-ray sources by virtue of the great increase in the total exposure that they will provide. Their complementary measurements determine source positions to several minutes of arc, low-resolution spectra from 1 to 60 keV, and variability over time scales from seconds to months.

A second generation of satellite x-ray experiments, now being prepared for 1973–1974 launches, will incorporate many advances in knowledge and technique gained from recent rocket and balloon experiments, in which new methods have been developed and tested for precise position determination, ultrasoft x-ray detection, high-resolution spectrum scanning, and background suppression. These methods are also suitable for much larger satellites such as the High-Energy Astronomical Observatory-A (HEAO-A). Detectors with sensitive areas as large as 10 m^2 or more could achieve a tenfold increase in the sensitivity of an all-sky survey with 1° resolution.

Long-range planning must be concerned now with the third generation of satellite x-ray astronomy experiments in which the emphasis will shift from general surveys to precise position determinations and the detailed examination and measurement of individual x-ray sources through high-resolution studies of their structure, spectra, and polarization. Of particular interest for cosmology is the high-resolution study of the x-ray background, with the aim of determining whether it is a superposition of many discrete extragalactic sources or is truly diffuse and therefore of intergalactic origin. Such experiments must employ grazing-incidence reflection x-ray telescopes to form high-resolution (~1 sec of arc) x-ray images, which can be electronically detected and corrected for residual spacecraft orientation error. The telescope will incorporate auxiliary instruments for high-resolution Bragg reflection spectrometry and polarization analysis. Small grazing-incidence reflection x-ray telescopes have already been used with outstanding success in solar observations that revealed the detailed structure of the x-ray emitting regions around flares. A 12-ft version will be used in the Apollo Telescope Mount-A (ATM-A) solar astronomy mission. A telescope of 20-ft focal length or greater, with an effective collecting area of 1000 cm^2, is within the current state of the art and could be placed in orbit by 1975 as a part of a large x-ray astronomy facility in NASA's projected HEAO series. Because of its high angular resolution, such a facility will have a sensitivity for the detection of faint sources that is quantum limited to about 10^{-7} cm^{-2} sec^{-1}. It could analyze the structure and polarization of extended objects and examine spectra for evidence of emission lines and absorption edges due to the elements from calcium to carbon. Perhaps most important of all, it could achieve position determinations to an accuracy of several seconds of arc, so that definite optical identifications could be achieved on even very faint objects.

New major satellite efforts are in order for the study of gamma rays. The successful low-resolution survey experiment of the Orbiting Space Observatory-3 (OSO-3), which was launched in 1967, was followed in 1972 by a high-resolution gamma-ray spark-chamber experiment on SAS-B. This experiment should delineate in considerable detail the gamma-ray distribution from the galaxy and throw new light on the origin of the extragalactic gamma rays. This instrument, however, is small for the job that needs to be done. The projected HEAO-A spacecraft would be well suited to a sufficiently large-scale gamma-ray experiment to achieve a significant improvement in sensitivity and possibly to permit the detection of extragalactic sources.

The central importance of continued balloon and rocket research to the vigorous growth of high-energy astronomy merits emphasis. Even after the first major x-ray astronomy experiments were orbited in 1970 and 1971, balloons and rockets continued to provide relatively quick and cheap means to explore new scientific and technical ideas, stimulate the development and testing of new instruments and methods, and train students in situations in which they could assume the major responsibility for carrying through entire experiments in high-energy space astronomy. This last function serves a wider need than that of providing competent scientists for future space astronomy. Successful balloon and rocket experiments generally require sophisticated and remote instruments, imaginatively designed for what are frequently speculative scientific purposes. They also require complex data processing and analysis, involving the wise use of computers. These challenges develop experimental skills that can be applied in many other fields of exploratory science.

There still remain large regions of the high-energy photon spectrum where the technical problems of achieving adequate detection capabilities have not been satisfactorily solved. One of these is the important region from 0.5 to 30 MeV, where extragalactic nuclear gamma-ray lines might be found and where the shape of the extragalactic spectrum could have important implications for cosmology. Although significant exploratory results have been achieved, the fundamental problem of obtaining a good signal-to-background performance from a directional detector has not been solved. Another energy region that may require much development before positive results can be achieved is that above 10 GeV, where air-shower techniques appear to offer the only promising line of approach. Here again, exploratory work has defined the character of the problem, but no adequate solution is clearly in sight. It is important that ways be found to support worthy new exploratory efforts in these regions of the high-energy spectrum.

Although present commitments and future plans promise a fruitful program of high-energy astronomy observations during the next decade, support for the critically important coordinated optical observations is inadequate. Rough distance indications can be obtained for x-ray sources in our Milky Way by purely x-ray techniques, but the greatest advances in the understanding of x-ray sources have occurred when they have been optically identified and studied with the precision tools of optical astronomy. Sco X-1, Cyg X-2, the Crab nebula and pulsar, the extragalactic sources M87, NGC4151, and 3C 273 are cases in point. As in the earlier experiences of radio astronomy, observations in the x-ray region of the spectrum provide the indispensable basis for new data revealing new types of phenomena and of objects. Only when the stream of information carried by optical photons is tapped with the powerful methods of optical astronomy can the full benefit of the x-ray observations be realized. Thus a balanced program in high-energy astronomy should provide adequate support for the related optical studies.

The optical counterparts of x-ray sources need to be identified by comparison of x-ray source positions with objects on photographs exposed to the faintest possible limiting magnitudes. Next, detailed studies of the spectrum must be undertaken. Often x-ray sources have proved to be optically variable, so it is important to monitor the optical variations for correlation with the x-ray data. These requirements point to the need for additional observing time on intermediate-size (2–4-m) optical telescopes that can be devoted to the monitoring and study of variable x-ray sources.

Other types of interesting objects are known to be variable, including supernovae, quasars, and galactic nuclei. The evidence on these objects from radio astronomy points to injection of fast particles and to variability at all wavelengths. Thus there is a need to monitor this wider class of sources as well.

Progress in this field can come by applying the most modern techniques of digitized recording of optical photons, including information about direction, wavelength, and polarization, at the focus of intermediate-size telescopes of modest cost. Such techniques can be applied to existing telescopes or, if the number available for this purpose is not adequate, to one or more new intermediate-size telescopes.

The instruments now under construction and planned for x- and gamma-ray astronomy from space portend a rich harvest of results important for relativistic astrophysics and cosmology. For example, detailed studies of the x-ray emission from neutron stars are possible, and large numbers of extragalactic sources will be found and identified. Detailed study of the x-ray and the gamma-ray background, which is of great importance to cosmology, will be accomplished. Consequently, this Panel emphatically supports a vigorous program of high-energy astronomy and makes specific recommendations to that end to the Physics Survey Committee.

Recommendation 15: High-Energy Astronomical Observatory. X-ray astronomy offers an opportunity to detect relativistic particles at their point of origin by means of their synchrotron and inverse Compton emission. The origin of enormous quantities of these particles in galactic sources such as the Crab pulsar and extragalactic sources such as radio galaxies and quasars is believed to be connected with violent events. X-ray observations may be decisive in disclosing the nature of these events.

The diffuse background at x- and gamma-ray wavelengths is apparently cosmological in origin. Future observations of its spectrum and angular structure could disclose whether it is truly diffuse, therefore intergalactic in origin, or perhaps the result of the addition of a large number of powerful x-ray sources.

The High-Energy Astronomical Observatory (HEAO) proposed by NASA would permit decisive contributions to the study of galactic and extragalactic discrete x-ray sources and of the diffuse x-ray background by achieving large increases in collecting area, angular resolution, and spectral resolution. The usefulness of the HEAO would be enhanced by a guest investigator program similar to the successful programs operating with other scientific satellites. Continued rocket and balloon research in high-energy astronomy will be necessary even after the HEAO becomes operational, in order to try out both scientific and technological innovations and to train students. *We therefore endorse the program for the development of a High-Energy Astronomy Observatory, and we recommend that an early start be made on a mission with a large grazing-incidence telescope capable of high angular and spectral resolution. We also recommend that balloon and rocket research in high-energy astronomy be continued at a reasonable level.*

Recommendation 16: Gamma-Ray Detectors. The gamma-ray region from 0.5 to 30 MeV, in which extragalactic nuclear gamma-ray lines may be found, is of particular interest to cosmology because of the light it will throw on explosive nucleosynthesis of the heavy elements in distant galaxies, the rate of which depends on the cosmological model. *We therefore endorse efforts to improve detectors in the range from 0.5 to 30 MeV and recommend that the best available instruments in this energy range be incorporated in High-Energy Astronomy Observatory payloads.*

Recommendation 4: Monitoring Variable Optical Objects. Several types of optical object of interest to relativistic astrophysics, including x-ray sources, supernovae, quasars, and galactic nuclei, are variable in the optical wavelength range. The time dependence, spectral characteristics, and polarization of these variations can yield significant physical information. To monitor such objects takes a significant fraction of the observing time on an intermediate-size telescope. *We therefore recommend that a number of intermediate-size telescopes be made available for substantial periods for such monitoring activities and that they be instrumented with detectors and data-handling devices adequate for precise and rapid data recording.*

5.5 COSMIC-RAY PHYSICS

Cosmic rays provide a probe of distant conditions and events that is only beginning to be exploited in astrophysics. Cosmic rays represent a sample of matter from the sun, distant stars, and perhaps even galaxies. Studies of the fast particles from the sun show that the relative abundances of most nuclei are approximately preserved during acceleration and transit from the sun. The chemical composition of the sun can be determined from the fast nuclei ejected during solar eruptions. The solar abundance of neon has been so determined. Presumably the galactic cosmic rays represent a sample of matter from pulsars, supernovae, and distant flare stars. Studies of the abundance of the various nuclei in cosmic rays at energies below a few GeV/nucleon (possible only with instruments carried on spacecraft) show, for instance, that the ratio of even to odd atomic numbers is large—of the order of ten—indicating that there is equilibrium of the nuclei in a condensed hot state ruling out rapid synthesis of these nuclei by collisions shattering heavier nuclei.

Cosmic-ray nuclei have been observed at energies of only a few MeV/nucleon; the persistence of heavy nuclei (Fe, Ni) at these low energies indicates that they may come from nearby sources, with an overabundance of heavy nuclei. The tracks of very heavy nuclei, such as Pb and U, in plastic detectors implies an enormous overabundance of very heavy nuclei accelerated from their sources.

Cosmic rays originate in energetic events. A few are produced in eruptions on the sun. The galactic cosmic rays seemingly are produced in supernovae, and supernova remnants (the pulsars). The very-low-energy cosmic rays must be produced nearby, perhaps in flare stars. On the other hand, the very energetic cosmic rays (10^{18} eV and above) are not confined to the galaxy and are presumably of extragalactic origin. It would be exceedingly interesting to know the nuclear composition of such very-high-energy cosmic rays, representing matter from distant galaxies.

Cosmic rays exert a significant pressure in space. They inflate the gas of the disk of our galaxy, preventing its collapse. The passage of cosmic rays through the interstellar gas ionizes and heats the gas, causing the gas to move closely with the magnetic field. The magnetic field, the cosmic-ray pressure, and the heating contribute to the instability of the gas. The general instability is the basic cause of the dumping of interstellar gas into clouds, where star formation apparently occurs.

Cosmic rays are deflected and modulated in the solar system by the magnetic fields carried in the solar wind. Observations of the variation of the cosmic-ray intensity at the surface of the earth provided some of the first clues (20 years ago) to the conditions in interplanetary space. With very-high-altitude balloons and space vehicles it has been possible to develop instrumentation and techniques to explore the subtleties of the cosmic-ray nuclear abundances and energy distribution over the entire lower end of the cosmic-ray spectrum. With such detailed information, progress in the field is rapid.

One outstanding question is the degree to which the solar wind and its magnetic fields reduce the cosmic-ray intensity throughout the inner solar system. Until we send suitable spacecraft instrumentation to the orbit of Jupiter (5 A.U.) and beyond, we will not know the true intensity of cosmic rays in the galaxy, particularly at the low-energy end of the spectrum, where they are much reduced by the wind, nor their integrated energy.

In summary, cosmic-ray physics is a major method of studying high-energy particles accelerated in violent cosmic events and the only method that uses direct samples of matter from the cosmos. The information gained by this method on the total energy density, lifetime, chemical and isotopic composition, and energy spectrum of various types of high-energy particles yields important insights into the nature of the sources where they were accelerated, particularly when combined with data derived from the electromagnetic emission by such events. Further experiments are required to verify the nature, frequency, and spatial distribution of source events, including extragalactic events that may be the source of the most energetic cosmic rays.

This Panel has not made a study of the experimental methods used in this area and therefore has refrained from making recommendations. However, it wishes to call attention to a critical situation developing in this field and to the response of our sister panels to it. In 1969, the Astronomy Missions Board in its report to NASA* recommended a vigorous program in particle and fields astronomy which has not been strongly implemented. Indeed, in a letter to the National Academy of Sciences Physics Survey Committee dated May 21, 1970, the NASA Associate Administrator pointed out that space physics (which includes cosmic-ray physics) is being deliberately de-emphasized in NASA to encourage other programs. This letter requested the Physics Survey to treat this subject explicitly.

The Panel on Space and Planetary Physics, in its report to the Physics Survey Committee† has done precisely that. It calls attention to the fact that the 1970 Space Science Board study, *Priorities for Space Research 1971–1980,*‡ protested strongly against the rapidity of deflation of funding in space physics. This is significant because the group responsible for this report was broadly representative of the entire space-science community not just space physics. The Panel on Earth and Planetary Physics concludes its discussion with this statement, "We recommend that the support proposed for space physics in the 1970 report on priorities for space research be regarded as minimal."

The report of the Panel on Space Astronomy of the Astronomy Survey Committee takes a similar view. It recommends this program:

(a) A deep-space probe to a distance of the order of 30 A.U. Such a probe would measure galactic cosmic rays directly, without the disturbing modulation by the solar wind.

* *A Long-Range Program in Space Astronomy* (NASA, Washington, D.C., 1969).
† Panel on Earth and Planetary Physics, Physics Survey Committee, this volume p. 849.
‡ Space Science Board, *Priorities for Space Research 1971–1980* (National Academy of Sciences, Washington, D.C., 1971).

(b) Pioneer F and G Jupiter flyby and the High-Energy Astronomical Observatory programs. These flights will permit a variety of cosmic-ray observations.

(c) A series of high-bit-rate interplanetary monitors. These spacecraft would particularly study the effects of the solar wind.

In addition, the Astronomy Missions Board called attention to the cost savings that can be effected by including space-physics experiments on planetary probes, in these words: "Indeed, it is absolutely essential to combine the interplanetary observations in cruise modes to the planets if costs during the next six to eight years are to be kept within manageable proportions." *After considering these matters, the Panel wishes to direct particular attention to the foregoing recommendations.*

5.6 SOLAR NEUTRINO ASTRONOMY

The Brookhaven Solar Neutrino Observatory, located deep in the Homestake Gold Mine in South Dakota, is continually improving its measurements of the neutrino flux from the sun; already the indications are quite exciting. The upper limit to the solar flux, which has not yet been positively detected, is considerably smaller than the original expectations generated by theoretical astrophysics. Continued searching will either increase this discrepancy or produce a positive detection. There are good reasons for this search for solar neutrinos: their detection will provide the most direct test of the hypothesis that the sun is generating thermonuclear power. The value of the flux will provide a stringent test of mathematical models of stars, on which much of our interpretation of the entire universe is based. It should be remembered that it takes over ten million years for energy in photons to reach us from the center of the sun but only 8 minutes for a neutrino. Hence the neutrino flux is the only sure indication of the present condition of the interior. This measurement provides an independent check on the physics we have been using to understand the universe. Essential for maximum results would be the development of practical schemes for detecting the lower-energy solar neutrinos from ^{7}Be, ^{13}N, ^{15}O, and the proton–proton reaction.

It should be remembered that the Brookhaven experiment is sensitive only to the total neutrino flux, from whatever sources. Because of its proximity, the sun is expected to be the dominant source. In principle, this supposition can be tested by observing an annual variation, expected to be 13 percent because of the radial motion of the earth in its elliptical orbit. Unfortunately, a much larger detector would be required for this purpose.

A related experiment is the laboratory measurement of neutrino–electron scattering. Such an experiment would measure the strength of the unknown electron–neutrino interaction, a value that plays a major role in the theory of advanced stellar evolution and in the possible coupling of neutrinos to the leptons and photons during the expansion of the fireball in big-bang cosmology.

Recommendation 10: Neutrino Astronomy. The attempt to detect solar neutrinos is critically important for all astrophysics. It is particularly so for relativistic astrophysics because of its implications for the theory of stellar evolution, the helium content of the sun, and the possible variation of the gravitational constant. These problems are closely related to the determination of the age of the galaxy, the problem of the formation of helium in cosmological models, and the choice between rival theories of relativity, all of which are critical for relativistic cosmology. *We therefore recommend that attempts to detect solar neutrinos be supported adequately until decisive results are achieved.*

5.7 GRAVITATIONAL RADIATION EXPERIMENTS

The detection of gravitational waves has been claimed recently on the basis of an extended series of experiments. This discovery would be of extraordinary significance and should be checked by continuation and elaboration of the original work and by other independent investigations. If it is confirmed, the result will not only be of fundamental significance for physics but also will imply the existence of totally unimagined relativistic astrophysical objects in our galaxy, the study of which probably will be crucial to understanding the structure and evolution of galaxies and could have major implications for stellar evolution and cosmology. Efforts to verify the detection of gravitational waves and the astronomical study of their sources should receive the highest priority, particularly since such work can be accomplished with modest expenditures.

The detection of gravitational waves bears directly on the question of whether there is any such thing as a "gravitational field," which can act as an independent entity. All actively pursued gravity theories deal with the concept of a gravitational field, so the mere existence of gravitational waves does not exclude any of these theories. (However, detailed properties of the waves can discriminate among competing theories, as discussed in Chapter 7.) Thus this fundamental field hypothesis has been generally accepted without observational support. Such credulity among scientists occurs only in relation to the deepest and most fundamental hypotheses for which they lack the facility to think differently in a comparably detailed and consistent way. In the nineteenth century a similar attitude led to general acceptance of the ether and atoms decades before the experiments that abolished the ether and confirmed the atom.

The basic style of all physics so far in the twentieth century has been set by the field concept, which arose in electromagnetic theory to replace the vanquished ether. This idea has been so overwhelmingly convicing, when tested in experimental and industrial applications, that scientists have tried to package every other known fundamental domain of physics in the same mold. Field theory is incontrovertibly successful in the case of the electromagnetic field. Application of field concepts to particle physics has been successful in many respects, but there are still many unresolved problems. Confirmation of the gravitational wave experiments would show that this concept is suitable for at least one of the other areas—that of gravitational phenomena—in which it is customarily employed.

The astrophysical implications of the gravitational wave experiments are profound and make it impossible, with any straightforward interpretation, to accept the initial observations without extensive confirmation. This situation is true even for resilient minds already stretched by the preposterous demands that radio galaxies, quasars, pulsars, and x-ray, gamma-ray, and infrared sources make on the astrophysical imagination. The gravitational wave observations could be a manifestation of some as yet unearthed subtlety. Otherwise, one relatively conservative interpretation seems to be to postulate that straightforward theory underestimates the sensitivity of the gravity antennas by several orders of magnitude, so that the emitters then could be only normally exotic by the standards of the past remarkable decade. Another interpretation, using the expected sensitivity, demands that our galaxy have but a small fraction of its original mass, with the bulk of it having been converted into gravitational radiation by a process of nearly perfect efficiency. Most of the energy in the universe might then be gravitational waves from similar galaxies. All models of sources for gravitational waves at the current receiver frequency require that masses of the order of one solar mass move at nearly the velocity of light and change their velocity by the same amount every millisecond in each brief burst of activity. Our curiosity to know whether these ideas must be faced seriously is intense and can be satisfied only by further experiments.

Recommendation 11: Gravitational Radiation Experiments. Recent experiments suggest that an enormous flux of gravitational waves could be present in space. Confirmation of the detection of such waves would constitute a crucial test of fundamental assumptions underlying the theory of gravitation. A flux of a magnitude even approaching the reported one would have extraordinary implications for astrophysical processes involving relativistic motions of astronomical objects. *We therefore recommend that experiments to detect gravitational waves and the study of their astronomical sources be fully supported.*

5.8 THEORETICAL STUDIES

Enormous effort has been devoted to the study of the theory of general relativity. From it, mathematical physicists have derived the extension of Newton's laws of gravitation to relativistic speeds, the interaction between point masses, the properties of gravitational waves, and models of stationary, rotating, and expanding masses that can represent stars, galaxies, or the universe. Yet the theory remains a difficult mathematical problem. Although it can be written in a single line, it embraces many nonlinear partial differential equations often having singularities of obscure origin.

The emphasis in the years ahead will be on the application to observable phenomena, such as the behavior of the universe at large red shifts and the rotational and vibrational modes of neutron stars. The flow of information to the theorist only a few years ago consisted of a few red shifts of distant galaxies. Now he must consider the isotropy and spectrum of the microwave background; the number counts of radio sources; the generation of large amounts of x-ray, optical, infrared, and radio power in galaxies and quasars, and the acceleration of fast particles therein; the structure of rotating, magnetized neutron stars; and the emission of gravitational waves by asymmetrically collapsing massive objects. This new information forces the theorist to be relevant to actual physical objects whose properties are constrained by the observations. Solutions of the equations having a high degree of symmetry (which permits mathematical rigor) will be of less interest in the future than solutions having less symmetry, which are obtained only on a computer but are nevertheless more like the real world.

Because of this trend, we can anticipate a rapid growth in the theoretical application of general relativity to the astronomical universe. The object of the work will be twofold. First, because general relativity itself is still not beyond doubt, detailed model calculations will be undertaken for the universe and relativistic objects, such as massive stars, to verify whether relativistic models will fit the data. If and when the correct theory is finally established by comparison with the data, such models will serve as an analytical tool to relate the observations to the basic phenomena. A historical prototype of this activity is the study of stellar structure based on the Newtonian theory of gravity and the quantum theory of the atom. Machine calculations based on straightforward but complex equations make it possible to infer the age, mass, composition, and internal temperature of a star from its external characteristics, such as luminosity and surface temperature. In the same way, we can ultimately hope to know the internal structure of a quasar, using the equations provided by general relativity.

One of the special problems encountered in this effort is that of singularities. Already there are available quite general theorems that prove that singularities must occur in a broad class of general relativistic solutions. This situation is almost unprecedented in classical physics, and the meaning is still obscure. As an example, we mention that in big-bang cosmology the whole universe emerges from a singularity in which density and temperature are infinite. On the one hand, the classification and prediction of these singularities will demand the efforts of mathematicians. On the other, singularities will call for the scrutiny of theoretical physicists, who no doubt will remain skeptical of any physical theory that predicts them. Such scrutiny might finally yield a modification of the theory of general relativity—for example, its quantization—which becomes important under extreme conditions and which may prevent the system from becoming truly singular.

As an example of a definite theoretical problem of observational interest that requires penetrating analysis, consider the collapse of a dead star whose mass is greater than that of a neutron star in equilibrium. Such an object must collapse at nearly the speed of light, each part interacting with every other according to the nonlinear field equations. Even if the configuration is relatively spherical at the start, rotation will ensure that it becomes less symmetric as the collapse proceeds. Moreover, it is likely that the system will be unstable, as gravitational energy can be released in various deformations of the surface or in subcollapses of internal parts. The whole object will be radiating such intense gravitational waves that precise calculation of them will be necessary to evaluate the radiation reaction, which affects the collapse itself. This calculation will require joint solution for the variables both inside and outside the star. As the object collapses, singularities develop, the dimensionality of which depends on the symmetries maintained in the collapse.

The solution to such a problem will require the services of mathematicians and physicists of the highest intellectual caliber. They should have ready access to computers of the greatest possible memory capacity and should work in close association with specialists in astrophysics, nuclear physics, plasma physics, and the like so that the relevant physical phenomena can be included as necessary.

Recommendation 12: Theoretical Studies. Application of the equations of general relativity to observable objects will be important to verify the correctness of the theory and to interpret correctly the basic phenomena that can be inferred only indirectly from observations. This activity will require the efforts of mathematicians and physicists of the highest intellectual caliber, together with the judicious use of the most powerful computers available. *We therefore recommend that individuals and groups doing outstanding theoretical work in astrophysics and relativity be adequately supported, and that the most powerful computers be made available to them.*

5.9 INSTITUTIONAL ARRANGEMENTS

In general, astrophysics flourishes best when there is good contact between physicists and astronomers and between observers or experimentalists and theorists. To any problem in astrophysics, the physicist brings his knowledge of the basic laws, which is vital to the proper understanding of the astronomical phenomenon, while the astronomer brings his knowledge of the astronomical context, important if a relevant and not misleading interpretation is to be found. The theorist will bring to the problem a knowledge of relevant mathematical technique and interest in the phenomenon as a manifestation of a more general class, while the observer or experimentalist is expert in experimental technique, aware of the limitations of his data, and vitally interested in tying together the diverse observational phenomena of which he is constantly aware.

Often we find a physicist-astronomer joined in one person; similarly, some astrophysicists are equally skilled in theory and observation. More often, however, we are dealing with different people, and in that case conscious steps must be taken to promote good communication. Virtually all departments of physics and astronomy place substantial emphasis on both theory and experiment. To encourage interaction between physics and astronomy, some universities have combined the two departments into one. In other universities this action has not been taken, but efforts are made, such as joint seminars and degree programs, to encourage communication between the separate departments.

How does relativistic astrophysics fit into this framework? It seems that the most productive institutions are those in which the interrelationships just described are most lively. So far there seem to be three discernible groups of workers in the field. First, there are those working on the theory of general relativity, who usually have a background in mathematical physics and who often work in mathematics or physics departments. Then, there is a small number working in physics departments on the experimental verification of the theory. Finally, there is a large and diverse group of astrophysicists, both theoretical and observational, who are studying a great variety of astronomical phenomena that have at least some relativistic aspects. These individuals can be at home in both physics and astronomy departments. Each institution must experiment with the mix of these groups and the departmental arrangements that lead to the most effective programs. It may be helpful to institutions not yet active in the field to know that the mutual stimulation provided by the three groups is a benefit that can be realized by making at least two or three positions available.

We have alluded several times previously to the fact that the instruments of importance to cosmology tend to be the largest of those employed by astrophysicists and astronomers because of the faintness of the objects studied. For financial reasons, many of these instruments will be built at national observatories, where they are accessible to the entire community, or at least at private institutions willing to share the facilities with outside users. The concentration of large instruments at a few institutions places a special responsibility on the management of these institutions to attempt to ensure that significant work on such long-range problems as cosmology takes place in spite of the many competing demands for observing time. We have already discussed this problem in Sections 5.1 and 5.2, and Recommendation 1 also is pertinent to it.

6 Impact on Other Branches of Science

6.1 OTHER BRANCHES OF PHYSICS

Astrophysics and relativity has considerable impact on other branches of physics. Its role in the testing of rival theories of relativity will be discussed in Chapter 7. Here we consider only the relation to the theory of elementary particles and to other better established subfields of physics.

Elementary-particle theory is strongly related to the theory of the early phases in the big-bang model of the universe. If current versions of the big-bang model are correct, then the universe once was vastly denser and hotter than now. To be able to solve the gravitational field equations, we need to know the equation of state of the matter and radiation present in the early universe, but our present understanding of particle physics is inadequate at temperatures above about 10^{12} K to 10^{13} K, at which strongly interacting particles are produced copiously in thermal equilibrium. One way to deal with this problem is to treat the matter as consisting of a number of species of highly relativistic free particles. If we take a fixed number of species (say, photons, gravitons, leptons, and nucleons), then the temperature of the universe is inversely proportional to its radius. If, on the other hand, we take as many species of particles as would exist in thermal equilibrium according to a currently fashionable model of strong interactions (the Veneziano model), then the temperature varies much more slowly. Probably no free-particle model makes sense; and to understand the early universe, elementary-particle theorists will have to leave the familiar conceptual framework of S-matrix theory and venture into the unknown territory of relativistic many-body physics.

One by-product of a realistic model of the early universe might be a clue to the existence of the hypothetical fundamental particles of strong-interaction physics—the quarks. Using a crude model, it has been estimated that if quarks are real, then enough should be left over from the hot early universe to make their current abundance about equal to that of gold atoms. Needless to say, quarks are a good deal rarer on earth than gold, and it would be important to know whether this absence really means that quarks do not exist.

Finally, we hope to learn more about elementary particles from the study of specific astronomical objects such as neutron stars and quasars. For example, the present theoretical uncertainty in the electron–neutrino interaction might be removed by studies of stellar evolution, in which, as stars approach the neutron-star phase, the annihilation of electron–positron pairs into neutrinos that can escape the star probably plays an important role. Neutron stars are fairly well described by nuclear physics, but there could be marginal effects depending on unknown physics. The strange phenomena associated with quasars offer hope of discovering even more unusual conditions of density and temperature that are not consistent with terrestrial physics. It has often been suggested that conservation of baryon number (a rigorous law of terrestrial physics) might be violated, perhaps by creation of matter as required by steady-state cosmology. The discovery of baryon number nonconservation would remove a long-standing puzzle: Why is there not an electromagnetic field coupled to baryon number as there is to charge, another conserved quantity? At the same time, of course, such a discovery would have profound implications for the origin of matter and would provide a possible explanation for the enormous energy output of quasars.

The impact on other branches of physics is pervasive. One could argue that Newton had to invent calculus and modern mechanics to solve an astronomical problem, and that they have subsequently found extraordinarily wide application in physics. The theory of atomic spectra was constantly challenged by the study of the sun and stars. Thermonuclear energy production was studied first in connection with the energy source for the sun and other stars and subsequently found spectacular terrestrial application. No doubt future developments in astrophysics will have equally unexpected and major impacts on physics.

Branches of physics that are supposed to be fully understood in principle have to be applied in a detailed and concrete way to situations that differ greatly from usual laboratory conditions. Physicists have thought for a long time that they understood thermodynamics and deviations from thermal equilibrium. Interstellar space is further out of thermal equilibrium than any laboratory apparatus—one finds both cosmic rays with an effective temperature of about 10^{14} deg and dust grains with a temperature of about 10 deg occupying the same region. One consequence of these deviations from equilibrium is maser action in interstellar space. Chronologically, masers were invented in the laboratory before they were discovered in the interstellar gas, but the problems posed by interstellar conditions stimulated research that is leading to an understanding of the relevant cooperative phenomena.

One aspect of this impact through application is cross-fertilization of different branches of physics. In discussing radiation mechanisms for pulsars, one has to employ both relativistic mechanics and plasma theory in detail; in calculating nuclear reaction rates in very dense stars, one needs solid-state techniques for dealing with zero-point vibration modes, as well as nuclear dynamics and the like. One characteristic flavor lies in the painstaking, matter-of-fact application of fundamental physics to the most bizarre and fanciful conditions in which natural objects find themselves. In particle physics the distinction between simple and composite particles is argued on a very esoteric level; for massive neutron stars this distinction can mean the presence or absence of Fermi pressure and, hence, the difference between stability and collapse. The fundamental study of strange particles is helped, at least indirectly, by their importance in engineering-type calculations for neutron stars.

On another level, cosmology raises questions about the division (always assumed to be valid) between the local laws of physics (which regulate what must be) and the actual properties of the universe (which govern what actually is). Laboratory physics is based on approximate symmetries such as charge-conjugation invariance (which states that the rates of particle reactions are nearly equal to those for the corresponding antiparticles). The universe probably does not embody these symmetries in its initial conditions, as matter appears to be much more abundant than antimatter. Is it possible that an interaction between the whole universe and local phenomena keeps the symmetry laws from being perfect? As yet there is no theory of this interaction, but it is a possibility. Another example is electrodynamics, in which the equations indicate a perfect symmetry between advanced and retarded potentials. The fact that only the retarded potential is actually observed could be related in some way to interactions with all the particles in the universe, which are known to be expanding rather than contracting. If this were true, the results of laboratory experiments might be intimately tied to the present state of the universe.

6.2 OTHER BRANCHES OF ASTRONOMY

The astronomer's laboratory is the universe. Since 1929, when the universe was found to be in a state of expansion, cosmology has been the central subject in astronomy. In an evolving model, the early more condensed stages of the universe contained the starting conditions for the formations of galaxies, clusters of galaxies, and quasars and, in particular, the Milky Way galaxy in which we live. Therefore, even astronomers studying stars within the Milky Way need the results of research on cosmology, just as cosmologists need the results of stellar and galactic astronomy on distances, time scales, and chemical abundances.

The finite velocity of light allows the astronomer to observe directly objects in the earlier stages of the universe. The light of such objects has traveled for many billions of years before it finally reaches us in the twentieth century. Even if the objects are as luminous as galaxies or quasars, at these enormous distances their signals on arrival at earth are weak. Large telescopes (x-ray, optical, infrared, or radio) are required to collect sufficient information in a reasonable time. As a result, observational advance in cosmology is primarily determined by the size and number of large telescopes. These telescopes are then available to study a host of other objects that may be intrinsically faint although relatively close by. In this way, construction of instruments for the study of cosmology tends to spur activity in all branches of astronomy.

6.3 EARTH SCIENCE

Although not always readily apparent, there are subtle relationships between the science of our solar system and relativistic astrophysics. Perhaps the most important connection is that between geochemistry and the origin of the elements. Much detailed information about abundance of the elements is determined from solid bodies in the solar system. The meteorites are preferred for this purpose, because they have apparently undergone much less chemical fractionation than has the earth–moon system. Because of their complexity, only a thorough understanding of their formation and chemical evolution will allow a confident interpretation of the conflicting abundance patterns revealed by them. For example, the abundance ratio $^{232}Th/^{238}U$ is quite variable in meteorites and lunar samples, and the correct ratio is important in determining the age of the galaxy. The choice between general relativity and the scalar–tensor theory of gravity may depend ultimately on an understanding of the geochemical fractionation among the elements U, Th, and Pb, because the two theories would assign different time scales to the galaxy.

Meteoritic and lunar sample studies offer a hope of finding stable superheavy nuclei. Some scientists have suggested that fission tracts due to such nuclei will be preserved in the meteorites and made visible by etching techniques. Definite information on the existence of superheavy nuclei would stimulate anew the study of the highly collapsed objects in which such nuclei might originate.

The question of the origin of life and the possibility of interstellar communication with life outside the solar system depend in part on the origin of planetary atmospheres, which in turn depends in subtle but significant ways on how planetary systems are born and how much young radioactivity is trapped in solid objects as they form—radioactivity capable of tipping the scales between molten and solid bodies. The rate of production of these same radioactivities plays a major role in the analysis of the age of our universe.

In addition to these specific examples, there are also philosophical relationships between cosmology and earth science. Broadly stated, earth science attempts to account for the origin of the earth and planets at a definite time in the past, their geological evolution through tectonic and atmospheric action, and the emergence and evolution of life on their surfaces. This evolutionary scheme probably is being re-enacted, with variations, countless times throughout the galaxy and the universe. Few scientists today believe that terrestrial life is unique, because astronomical research indicates that the building blocks of the universe—the stars and galaxies—are remarkably similar. Not only are there billions of stars virtually identical to the sun that presumably have planets similar to the earth, but also the relative abundances of chemical elements available for the extraordinary process we know as life seem to be virtually the same everywhere. This uniformity is simply a reflection of the uniformity of the universe as a whole. For example, carbon, a key element for life, is present everywhere, because galaxies were formed throughout the universe with similar properties. Therefore, the stars within them have similar masses and rotation rates and similar evolutionary histories. Hence the process in which three alpha particles in the interior of the star join to form carbon nuclei proceeds similarly everywhere.

Thus, evolution of the earth and of life should be viewed not as an isolated phenomenon but as one typical of a huge number of similar events scattered throughout the universe. Cosmology forms the giant canvas on which the evolution of the universe is painted, but life is given to the picture by those mysterious processes, occurring throughout the cosmos, in which matter ultimately evolves to consciousness, so that the universe becomes aware of itself.

7 Testing General Relativity

7.1 PHILOSOPHY

In considering astrophysics as a bridge between physics and astronomy, two aspects of the science merit special attention. One relates to situations in which a new phenomenon of particular interest to physicists is encountered in the astronomer's domain. In the other, relativistic gravitational phenomena are involved.

The unique importance of astronomical bodies for testing general relativity results from the great weakness of gravitation. The strength of the gravitational interaction between two elementary particles is roughly 10^{-40} of the electromagnetic forces that dominate laboratory physics. In other words, the amount of space curvature under general relativity is negligible over dimensions of space–time as small as a laboratory or the two years' duration of PhD research.

Tests of general relativity require bodies of astronomical size. If an astronomical body is as massive as the sun and as compact as a neutron star, the relativistic effects should be quite large. In fact, the binding energy is ~100 MeV/nucleon, larger than any other known force in nature. Throughout the solar system, relativistic effects are minuscule, but the possibilities of precision measurements in the solar system have led to the only two presently known positive tests of general relativity: the gravitational deflection of light and the relativistic rotation of Mercury's perihelion.

It is a mistake to think of these two relatively poor observations as providing our only basis for relativistic gravitational theory. More important are the null gravitational experiments and the various laboratory experiments on which we base our confidence in special relativity, the root-structure of general relativity. Of the null experiments, the spatial isotropy experiment and the modern version of the Eötvös experiment deserve particular notice. These are important in eliminating a large number of otherwise possible theories and in supporting the equivalence principle. Any acceptable gravitational theory must automatically yield a gravitational acceleration of small bodies substantially independent of composition (to less than one part in 10^{11}). It must be explicitly noted that similar observations involving very massive falling bodies are much poorer.

The isotropy experiment effectively eliminates any theory leading to significant anisotropic gravitational–inertial effects. Thus, a Lorentz-invariant tensor theory of gravitation can be eliminated unless its form permits the unification of the Minkowski metric tensor and the field tensor. Similarly, the various null experiments and other laboratory-based tests of special relativity impose limits on gravitational theories. Any acceptable relativistic theory of gravitation might be expected to yield special relativity over the small volumes of space–time required for laboratory physics, and such an acceptable theory must yield results for laboratory physics in agreement with the observations.

With the assumption that the gravitational theory shall have Lorentz-invariant (or special relativistic) roots, a formal machinery exists ready-made for the description of gravitation, that is, the Lorentz-invariant field theories already developed for the description of electromagnetism and particle physics. Gravitation requires one or more chargeless, massless boson fields for its description. Allowable fields have spin 0, 1, or 2. A theory using a neutrino field (spin ½) has certain technical difficulties. The spin-1 field (the analogue of the electromagnetic field) leads to a repulsive force and has other difficulties. Only the spin-zero (scalar) field, spin-two (tensor) field, or a mixture (scalar–tensor theory) remain for consideration.

The scalar field theory developed by Nördstrom in 1912 is satisfactory in most ways, but its predictions concerning the gravitational deflection of light and the relativistic perihelion rotation are not in agreement with observation. Both the tensor theory and the scalar–tensor theory are satisfactory, and it is not yet clearly established which of the two theories receives the most observational support. The scalar–tensor theory has certain interesting properties; for example, it provides insight into, and a means of calculating, the coupling constant of gravitation, but this theory is more complicated than the tensor theory, as it requires two fields. For this and other reasons few scientists favor the scalar–tensor theory on either philosophical or aesthetic grounds. But observations should be the primary concern, not aesthetics.

A modest goal for the future is to devise a test of gravitational theory sufficiently accurate and unambiguous to permit the exclusion of one or the other of these two theories. A more ambitious goal is to devise enough independent tests of the remaining theory to provide strong observational support for it, if sufficient funds for the tests were made available. If it should happen that neither of these two theories is tenable, the theoretical implications would be both serious and interesting.

7.2 EXPERIMENTAL TESTS USING THE SOLAR SYSTEM

It is a measure of Einstein's genius that in his first comprehensive paper on general relativity, his tensor theory of gravitation, he suggested all the positive tests of general relativity known until very recently. These are the gravitational red shift, the gravitational deflection of light, and the relativistic rotation of Mercury's perihelion. The gravitational red shift does not require the full machinery of general relativity for its discussion and is more properly considered to be a test of the equivalence principle. Thus it is more closely related to the null experiments than the other two tests, which investigate the particular form of the metric about the sun and distinguish between the two theories. Under the scalar–tensor theory the gravitational deflection of light should be $(1 - s) \times (1.75$ sec of arc) for a light ray passing close to the sun. Here, 1.75 sec of arc is Einstein's value for the deflection of a light ray grazing the sun's limb and $s = 1/(2\omega + 4)$ is the fraction of a body's weight due to the scalar field under the form of the theory for which Einstein's equations are formally satisfied, where $\omega \sim 5$ is the coupling constant of the scalar–tensor theory. If $\omega = 5$ (see below), $s = 0.07$, a measurable effect in the deflection of light.

The classical measurement of gravitational light deflection during solar eclipses has given rather poor results, but the techniques used probably could be substantially improved. In this case the importance of a single measurement is so great that a sustained effort to develop a special instrument and technique might be warranted. The great importance of this observation stems from the relative lack of ambiguity in its interpretation. The light deflection due to the solar corona should be much too small to be significant. To avoid the necessity for a total eclipse of the sun, an instrument is being developed capable of photoelectrically measuring star positions near the sun. An alternative approach is to use radio waves, which are less seriously affected by the glare from the sun. Using long-baseline interferometers, one can determine the positions of point radio sources near the sun, it is hoped with enough precision. Here the refractive effects of the solar corona can be important, requiring the use of two or more wavelengths or a sufficiently short wavelength. These approaches should be capable of distinguishing between the two gravitational theories. Experiments so far performed have yielded a value of $s = 0.06 \pm 0.06$, precision as yet insufficient to distinguish between $\omega = 0$ and $\omega = 5$.

An alternative approach makes use of the retardation of a radar wave passing the sun. This retardation is closely related to the deflection of the wave, as it is the retardation near the sun that wheels the wavefront about, changing its direction. This effect in radar returns from Mercury and Venus has been used to determine the related light deflection, obtaining the result $s = 0.1 \pm 0.2$, where the error is twice the standard deviation.

All three of the new techniques, as well as an improved technique based on a solar eclipse, seem to afford the precision necessary to discriminate between the two theories. Because of the relative lack of ambiguity in the interpretation of results, this type of observation shows the greatest promise of conclusively excluding one of the two competing theories.

The relativistic precession of Mercury's perihelion was known by the middle of the nineteenth century as an unexplained excess motion. At present, the analysis of the observations of Mercury give an excess motion of 42.3 ± 0.5 sec of arc/century, compared with Einstein's calculated relativistic motion of 43.3 sec of arc/century. Only the planetary perturbations are subtracted to obtain the above result. If the observed solar oblateness of 5×10^{-5} has been correctly interpreted as implying the existence of a solar gravitational quadrupole moment, an additional perturbation of 3.4 sec of arc/century must be subtracted, leaving 38.9 sec of arc/century for the relativistic effect. Under the scalar–tensor theory, the relativistic rotation of Mercury's perihelion is $(1 - \frac{4}{3}s) \times 43.3$ sec of arc/century. This expression would yield the observed value of the precession corrected for the quadrupole effect (38.9 sec of arc), if $s = 0.07$ or $\omega = 5$. Thus, scalar–tensor theory, with $\omega = 5$, is favored if the sun has the quadrupole moment indicated by the oblateness observations. The discrepancy with conventional general relativity imposed by the solar oblateness is in excess of three standard deviations, if there are no systematic errors. However, there has been considerable controversy about the significance of the solar oblateness.

Space science provides an unusual opportunity for observing and separating the relativistic and quadrupole effects. An artificial space probe moving about the sun in an elliptical orbit and carrying a radar transponder could provide such a new measurement of a relativistic perihelion rotation. Longevity of the space probe and a careful compensation of radiation pressure appear to be essential for a successful experiment of this type. An alternative and more romantic approach is to soft-land a transponder on one of the asteroids. Radiation pressure is too weak to affect significantly the motions of large asteroids. Both of these approaches would also permit the determination of "light deflection" using the retardation method if the transponder operates in the shorter microwave region of the spectrum.

Other relativistic effects should be measurable in the solar system, but they have not yet been observed. Among the interesting effects are the Thirring-Lense, or inertial drag, effect, which should appear near a rapidly rotating massive body, and the geodesic precession of a gyroscope, caused by translating it through a closed orbit in curved space. In principle these effects could be observed as a slow precession induced in a gyroscope. For such a gyroscope, a freely floating spinning top in a satellite orbiting the earth has been suggested. A satellite system to measure these effects is presently being designed.

Unfortunately, present accuracies do not permit a definite exclusion of one of the two competing theories, but this situation provides a challenge for the immediate future. It is hoped that the response to this challenge will be a gradually improving case for the correctness of the remaining theory. If neither theory should be correct, the situation would be even more interesting and could lead to an entirely new approach.

7.3 COSMOLOGICAL TESTS INVOLVING OBJECTS OUTSIDE THE SOLAR SYSTEM

Relativistic effects are minuscule in the solar system, but the close proximity of the sun and planets permits measurements sufficiently precise to be of interest. Outside the solar system compact systems for which relativistic effects are appreciable exist (neutron stars and possibly quasars). Here there is a possibility of investigating gravitational theory through gravitational radiation. If any of these compact bodies were to be a source of gravitational radiation, this radiation might conceivably be a source of information about the radiation mechanism as well as the radiating system. Gravitational radiation, whatever its source, is capable of providing a test distinguishing between the tensor and scalar–tensor gravitational theory. A spherical resonator can be excited in a radial mode by gravitational radiation under the scalar–tensor theory but not under the tensor theory.

Compact bodies are not the only possible sources of new gravitational tests. The enormous stretches of space and time available in the universe offer the possibility of observing relativistic effects even though the average space curvature is small. Cosmology is capable of providing interesting tests of gravitational theory, but this approach suffers from incomplete data and ambiguity in the interpretation of the available data. Several misconceptions exist concerning cosmology as a source of tests of general relativity. Thus, the existence of an expanding universe that appears to be reasonably isotropic and uniform in mass distribution, in the small part seen by us, is sometimes interpreted as support for general relativity. This view is based on the existence of expanding space solutions of the field equations of general relativity. Solutions of this type exist in both relativistic theories, the tensor theory and the scalar–tensor theory, but the dynamical equations that relate the deceleration parameter to the mass density do not require a relativistic calculation. Also, the initial conditions leading to the uniform and isotropic solution are as mysterious under relativity as under Newtonian cosmology. Although we currently do not understand the origin of this degree of order in the universe, we may some day be able to use the observations to test a theory of the origin of the order, if such a theory appears.

One particularly interesting but still unobserved aspect of relativistic cosmology concerns the relation between image size and distance. For relativistic cosmologies, a minimum angular size occurs at a certain distance that is similar for the tensor and scalar–tensor theories.

Under the tensor theory there are no locally observed cosmological effects. The enormous mass of the universe expanding away from us is without effect on our galaxy or solar system. There has been confusion about this point. Some publications mistakenly claim that the galaxy becomes entrained in the expanding space and expands slowly with time.

Under the scalar–tensor theory there is a possibility of observing locally the effects of the matter distribution. The scalar, being generated by the matter content of the universe, increases slowly with time. This growth results in a decrease of the gravitational "constant" with time. The weakening of gravitation with time carries a number of implications. In principle, these implications could provide information that would lead to the rejection of one of the two relativistic gravitational theories; in practice, this result is not yet possible. The following examples show the difficulty inherent in this type of gravitational test.

The present rate of decrease of the strength of gravitation expected under the scalar–tensor theory depends on ω and the present matter density. A rate of decrease of 10^{-11} part per year would be reasonable. This carries the following implications if the scalar–tensor theory is correct: Our galaxy and solar system should be expanding at the rate $\sim 10^{-11}$/year, and periods of planets and the moon should be increasing $\sim 2 \times 10^{-11}$/year as measured by atomic clocks.

Other differences between expectations under the tensor and scalar–

tensor theories depend on the effect of weakening gravitation on stellar luminosity. For solar-type stars this effect varies as the seventh power of the gravitational constant. This variation affects the red shift magnitude diagram slightly and the computed ages of population II stars substantially, making them smaller. It could affect the temperature of the earth in the remote past.

If we assume the correctness of cosmology based on the expansion of an initially hot fireball, expectations will differ under the two gravitational theories. Under the tensor theory, nearly independent of present mass density and neutron half-life, roughly 27 percent helium is formed in the fireball (assuming a ratio of neutrinos to nucleons $<10^7$). Under the scalar–tensor theory, if the scalar field contribution to the energy density in the fireball is sufficient, the early expansion rate might be sufficiently accelerated to stop helium formation. The initial fragmentation of the expanding fireball to form gas clouds (some of which may be represented by fossil remnants in the form of globular clusters) depends on gravitational theory. Under the tensor theory the expected mass of globular clusters can be almost an order of magnitude greater than under the scalar–tensor theory.

The techniques of space science also could be capable of yielding a gravitational test. The original motivation for placing a laser reflector on the moon was to test for the acceleration of the moon's motion expected under the two theories. Referred to a planetary ephemeris time scale, the acceleration due to tidal interactions is known over the past 200 years. Assuming that the tidal interaction has been constant for the past 200 years and will continue constant for the next ten years, we can separate the two effects and calculate the effect of weakening gravitation if this occurs. An alternative approach is provided by planetary radar. This technique is capable of detecting in a decade the gradual change of planetary periods, should changes occur. The lunar laser reflector is capable of still another type of test of gravitational theory. Under the scalar–tensor theory the gravitational acceleration of a body depends on the fractional contribution of gravitational self-energy to the body's mass. This contribution is appreciable for planets and differs for the earth and moon. This difference in acceleration leads to a displacement by less than a meter in the moon's orbit relative to the earth under the scalar-tensor theory.

Recommendation 17: Testing General Relativity. The choice between rival theories of gravitation cannot be made conclusively on the basis of present data. This choice is of fundamental physical significance. Moreover, work in relativistic astrophysics depends critically on this choice. *We therefore recommend that experiments using optical, radio, and radar* *methods to observe the deflection of electromagnetic waves by the sun, the retardation of such waves passing the sun, the precession of the perihelia and apsides of bodies orbiting the sun at various distances, and a possible lengthening of the orbital periods of such bodies be supported and emphasized within a well-balanced program of ground-based and space-based astronomy. In addition, the use of artificial satellites to detect the inertial drag and geodesic precession in earth orbit should be supported.*

8 Manpower, Funding, and Education

8.1 CHARACTERISTICS OF THE ASTROPHYSICS AND RELATIVITY SUBFIELD

Unlike most other subfields of physics and astronomy, astrophysics and relativity is not widely recognized as a distinct discipline. Therefore, manpower surveys and federal funding categories do not provide clearcut information on the statistics of the subfield. Consequently, we caution the reader that the statistics in this chapter are largely approximations.

The Panel attempted to identify a core group of scientists whose research is primarily in astrophysics and relativity, as defined in the first chapter of this report—the study of astronomical phenomena requiring general relativity for their interpretation. The Panel identified those federal funds that clearly are supporting the efforts of this group and attempted to estimate the additional funds that support the observational programs used less directly by this group, even though these funds have not been specified by the federal agencies as being for this purpose.

In this analysis we found that the observational programs are far more costly than the direct support of the core scientists involved. As we explain in the introduction to our recommendations, such expenditures can be justified in part by their application to the subfield alone, as it has high scientific interest, but the balance of justification must be provided by the many other possible applications of the instruments. In the conclusion to this chapter, this aspect of our recommendations is called to the attention of the two parent committees to which this report is addressed.

The Panel found that doctoral education for work in astrophysics and relativity is proceeding well in both astronomy and physics departments. With certain suggested improvements, there is no reason why this pattern should not continue.

8.2 MANPOWER

The National Register contains data assembled from American Institute of Physics (AIP) questionnaires sent simultaneously to both physicists and astronomers. From the list of specialties in the Register the Panel identified the five in Table VIII-1 as basic activities in astrophysics and relativity.

On the same basis, 634 PhD's in the 1970 Register Survey indicated definitely nonrelativistic astronomical categories such as binary stars and planets and satellites. In addition, some 150 of the 725 scientists in the Register who were identified with earth and space science worked in subfields such as solar and planetary physics, solar wind, and the sun and probably would be assigned to astronomy, for a total of 1041 PhD astronomers. If we take into account an 85 percent rate of return on the Register questionnaires, we arrive at the data in Table VIII-2. The total number of PhD astronomers estimated in this way, 1224, agrees fairly well with the number, 1256, established by the Astronomy Survey Committee through direct inquiries to 171 institutions known to be active in astronomy in 1970. [See *Astronomy and Astrophysics for the 1970's. Volume 1, Report of the Astronomy Survey Committee* (National Academy of Sciences, Washington, D.C., 1972).]

So far, the subfield of astrophysics and relativity has been considered as part of astronomy (because the rationale for funding relates principally to astronomical operations at the present time). In fact, most specialists in astrophysics and relativity obtained their PhD's in physics, as Table VIII-3 shows. In terms of training, the PhD population of the subfield is

TABLE VIII-1 Core Manpower in Astrophysics and Relativity, 1970[a]

Specialty	PhD's	Non-PhD's
Gravitational fields, gravitons	26	8
Cosmology	16	8
Galaxies	45	17
Quasars, pulsars, and x-ray sources	84	40
Relativity, gravitation	77	35
	248	108
Other[b]	9	51
Total respondents	257 (62%)	159 (38%)
Student respondents		80

[a] Data in the table are based on the National Register of Scientific and Technical Personnel.
[b] Respondents definitely in astrophysics and relativity, but for whom some items of Register data are missing.

TABLE VIII-2 Astrophysics and Relativity as a Portion of Astronomy PhD Manpower, 1970[a]

Subfield Identification	Number	Percent
Astronomy, categories other than astrophysics and relativity	745	61
Astronomy, part of earth and space physics	176	14
Astrophysics and relativity	303	25
Total astonomy PhD manpower	1224	100

[a] *Corrected for 85% rate of return.*

a mixture of physics and astronomy, as one would expect from the nature of the field. The number of physics PhD's engaged in the subfield (212) is 1.3 percent of all physics PhD's.

Some other characteristics of the 257 PhD's identified as part of the subfield are as follows:

1. Employment is heavily concentrated in the universities (73 percent) as compared with that of the overall population of PhD astronomers (63 percent) and PhD physicists (50 percent). Only 18 percent work in government laboratories or research centers, compared with 29 percent of the PhD astronomers. Only 16 percent work in industry or research centers as compared with 36 percent of all PhD physicists.
2. There is a heavy commitment of time to research and teaching (66 percent and 29 percent, respectively) as compared with the physics PhD population (37 percent and 50 percent, respectively).
3. The number of theoreticians is greater (51 percent) than in astronomy (28 percent) or physics (25 percent).
4. The PhD group is relatively young, with a median age of 34.7 years, compared with 35.0 for PhD astronomers and 37.4 for PhD physicists.
5. The PhD population is rapidly growing; between 1964 and 1970,

TABLE VIII-3 Field in Which PhD Was Obtained

	Field of Current Research		
Field of PhD	Astronomy as a Whole (%)	Physics as a Whole (%)	Astrophysics and Relativity (%)
Physics	36	80	70
Astronomy	57	5	25
Other	7	15	5

the number of PhD physicists increased by 60 percent, the number of PhD astronomers by 62 percent, and the number of PhD's in astrophysics and relativity by 300 percent (from 65 to 257 individuals).

The picture that emerges is of a rapidly growing group of relatively young university scientists, heavily committed to research, primarily theoretical. From personal observation, the Panel believes that a typical pattern is for a senior physicist to shift his research interest to this subfield and to encourage a similar move by a considerable number of graduate students and postdoctoral fellows who are eager for an opportunity to enter the subfield. Clearly, there does not appear to be a serious problem in the near future in obtaining trained manpower to exploit the many opportunities in this subfield, as at this time the number of PhD's in the subfield is only 1.5 percent of the total of 16,631 PhD's doing physics or astronomy. Large numbers of the remaining PhD's are in subfields from which movement into this one is relatively easy.

8.3 PATTERNS OF FUNDING

Certain federal agencies, the Department of Defense (DOD), the Atomic Energy Commission (AEC), and the National Science Foundation (NSF), supplied the Panel with data on support for astrophysics and relativity for the fiscal years 1965–1970. Most of the funds were expended in areas commensurate with the definition by specialties in Table VIII-1. The average of about \$4 million per year (\$4.3 million in fiscal year 1970) supports the work of a substantial fraction of the number of specialists in Table VIII-1. For example, probably most of the workers in gravitation and relativity (40 percent of the PhD's in Table VIII-1) are supported at less than \$25,000 per year, for a total of \$2.5 million per year. On the other hand, a few of the workers represented in Table VIII-1, for example, optical and ultraviolet space astronomers working on galaxies, radio astronomers working on quasars, and space astronomers working on x-ray sources, no doubt use much larger amounts of money in their research. It is difficult to decide exactly how much, because the agencies that supplied funding data did not attempt to include in their statistics that fraction of the support of extensive optical, radio, and space facilities used in studies of astrophysics and relativity. Moreover, no figures were available from the National Aeronautics and Space Administration (NASA). We have attempted to estimate figures to fill some of these gaps.

First, as Table VIII-2 shows, astrophysics and relativity manpower constitutes roughly 25 percent of all astronomers. This figure is corroborated at least roughly by the observation that about one third of the articles in the *Astrophysical Journal,* a leading periodical devoted to astronomy and astrophysics, are in astrophysics and relativity.

The Panel believes, from its knowledge of activities at well-known ground-based observatories such as Hale, Lick, Kitt Peak, the National Radio Astronomy Observatory, and Arecibo, that 25 percent is a reasonable estimate for the fraction of effort devoted to projects of interest to astrophysics and relativity. We, therefore, should assign to the subfield federal funding equal to about one fourth of the total ground-based astronomy program ($49 million in fiscal year 1970), or $12.3 million per year. Note that this is in addition to the $4.3 million per year identified by the agencies as direct support. It is believed that the latter is largely supporting programs in physics without large components in observational astronomy. Thus, even if we count only ground-based work, the expenditures in astrophysics and relativity are probably some $(12.3 + 4.3) \div 4.3 = 4$ times larger than the direct support of the core scientists in Table VIII-1.

When space-based research is included, the factor becomes even larger, although estimates are necessarily very imprecise. Of the current NASA astronomy program, we estimate that about 20 percent (including parts of OAO, OSO, Explorer, rockets, airplanes equipped to make infrared observations, and Supporting Research and Technology) is devoted to astrophysics and relativity. This amount includes ultraviolet observations of galaxies, study of x-ray sources, some cosmic radiation experiments, infrared observations of quasars, and the like. As the total cost of the 1970 NASA Astronomy Program (including management costs) is estimated at $220 million, some $44 million per year should be assigned to astrophysics and relativity. Funding data are summarized in Table VIII-4.

The figure of $60.6 million per year is obviously quite uncertain. It represents 23 percent of the total support of astronomy in fiscal year 1970 ($260 million) and is 14 times the direct support identified by the agencies. We believe that this "iceberg effect" of about one order of magni-

TABLE VIII-4 Federal Funding of Astrophysics and Relativity, Fiscal Year 1970

Nature of Support	Amount ($Million)
Direct support identified by DOD, AEC, and NSF	4.3
Ground-based observations (estimated 25% of total ground-based astronomy program)	12.3
Space-based observations (estimated 20% of total space-based astronomy program)	44.0
Total support, astrophysics and relativity	60.6

tude is a fair estimate of the real cost of carrying out the difficult observational and experimental programs that are vital to progress in this field.

When \$60.6 million is divided among 303 PhD's, the result is \$200,000 per year per PhD. Thus, astrophysics and relativity, in common with astronomy generally and with high-energy physics, is relatively expensive. Of course, as we see in Table VIII-4, this expenditure reflects the high cost of space experiments. Without the space component, which provides crucial observations of distant relativistic objects, the cost drops to \$54,000 per year per PhD, a figure corresponding more closely to the level of support in most physics subfields.

8.4 THE NATURE OF PhD PROGRAMS IN PHYSICS AND ASTRONOMY

Most workers in relativistic astrophysics hold a PhD in either physics or astronomy. Ideally, of course, each researcher should be thoroughly versed in both fields, but it would take such an extended time to achieve this that it is usually not possible within a reasonable PhD program. It is quite possible for physicists to acquire sufficient knowledge of astronomy after completing the PhD to work effectively in this subfield, but it is somewhat more difficult for an astronomy PhD to study the required physics later. Thus, although it is desirable for physics students interested in relativistic astrophysics to study some astronomy courses (for example, stellar evolution, stellar dynamics, galaxies and cosmology, radio astronomy, and high-energy astronomy), it is essential for astronomy students interested in this subfield to study relevant branches of physics. Those intending to specialize in theory should include advanced general relativity, together with study of the relevant mathematics, as well as advanced courses in one or more other areas of physics, such as quantum theory and statistical mechanics.

As we have pointed out in Chapter 5, pursuit of mathematical problems in general relativity will be directed increasingly toward relevant problems in astronomy, and, for that reason, it is helpful if students working on such problems associate with observers. Such activities have traditionally been carried out in physics departments; therefore, it will be necessary to arrange joint activities with astronomy departments, where the observers are.

Thesis work on experimental tests of relativity has been localized in a few places so far, usually in physics departments, where the style is divergent from much of current specialized research in particle or solid-state physics in that a typical experiment may draw on several applied disciplines, including optics, electronics, classical dynamics, cryogenics, and techniques of radio astronomy.

Traditionally, work on theoretical astrophysics has been centered in astronomy departments, but an increasing number of physics departments are active in this area. The physics student brings to a thesis in theoretical relativistic astrophysics a thorough training in fundamental physics, which is a solid advantage. On the other hand, it is helpful, to increase the relevance of any calculation performed, for the student to talk frequently with observational astronomers, who are usually located in astronomy departments.

Observational and experimental work on astronomical objects takes place in both physics and astronomy departments. Typically, an observer from the astronomy department uses traditional instruments (such as optical or radio telescopes) to observe a large number of objects of a certain class in order to analyze differences and similarities. A student in the physics department is more likely to build new apparatus to make an altogether different kind of observation—perhaps in a new wavelength range—and then apply the apparatus to one or two critical examples of a phenomenon. Both methods are necessary. It is clear that the activity in physics departments has been critically important in opening to investigation new wavelength ranges, which are the basis for much of the excitement in the subfield. There is no apparent reason why the design of new apparatus should not receive greater emphasis in astronomy departments.

8.5 RELATIVISTIC ASTROPHYSICS AS TRAINING FOR APPLIED SCIENCE

Today students are increasingly aware of the social implications of science. While they are drawn by the intellectual stimulation of a field like physics or astronomy, where elegantly simple laws govern extraordinarily complex phenomena, they also believe that physical science should be applied more vigorously to alleviate the problems that plague mankind.

Therefore, a major question is whether training in relativistic astrophysics can be useful for careers in applied science, should the student later make this choice. Astrophysics deals with the macroscopic world and as a result often spans much of physics that is relevant to applied problems. To understand theories of galactic or stellar formation and evolution, or recent models of quasars, pulsars, and the fireball, it is necessary to understand transport theory, plasma physics, magnetohydrodynamics, turbulence theory, nuclear physics, statistical mechanics, physical optics, classical electrodynamics, atomic physics, and the like. Such broad training in disciplines that can be applied to practical problems gives the student of astrophysics opportunities in such fields as weather prediction, aircraft design, nuclear-power development, and environmental control.

An advantage of relativistic astrophysics in particular is that it currently is focused on basic physical questions such as new forms of energy and the ultimate origin of matter; therefore, it can attract excellent young minds eager to attack basic questions. Having acquired a broad training in astrophysics, a student is then equipped to carry out intensive studies of specific astronomical phenomena or, if he so chooses, to use his knowledge to apply physics to human welfare.

8.6 IMPLICATIONS OF THE RECOMMENDATIONS FOR FUNDING

The recommendations of this report are addressed to both parent committees, but they fall naturally into those of primary concern to the Astronomy Survey Committee and those of primary concern to the Physics Survey Committee. In the estimates given in Table VIII-5, the recommendations have been grouped to emphasize this division of major interest and responsibility.

Estimated costs of the recommended programs are summarized in Table VIII-6. The fraction of each recommended program that will be devoted strictly to astrophysics and relativity has been roughly estimated so that these costs can be compared with the costs of the present program in astrophysics and relativity estimated earlier in this chapter. Again we remind the reader that these estimates are very uncertain.

The total program recommended in this report is estimated to cost \$1160 million over ten years, or \$116 million per year. Of this, the effort devoted to astrophysics and relativity is estimated to cost \$677.4 million, or \$67.7 million per year. Thus, the additional cost averaged over the years is about equal to the cost of the present program (Table VIII-4); therefore, it is equivalent to a program that averages about twice the present program over the next decade. This would imply an average growth rate of 14 percent per year. The Panel considers that such a rate of growth is commensurate with the scientific value of research in astrophysics and relativity.

8.7 IMPACT OF CONSTANT OR DECREASING LEVEL OF SUPPORT

Although charged to discuss this question, the Panel did not make a thorough study of it, so we confine ourselves to a few remarks. As ex-

TABLE VIII-5 Summary of Recommendations and Estimated Costs

Subject Matter of Recommendation	Cost, Including Ten Years of Operation ($Millions)
GROUP A[a]–AG[b]	
Recommendations	
Additional large optical telescopes	50
Monitoring variable optical objects (two 90-in. telescopes)	20
Electronic optical imaging	15
Large radio array	100
Large millimeter dish	20
Ground-based infrared telescopes	25
Theoretical studies	30
TOTAL	260
GROUP A[a]–AS[c]	
Recommendations	
Diffraction-limited optical space telescope[d]	300
Aircraft, balloons, and rockets[e]	125
High-Energy Astronomical Observatory	400
TOTAL	825
GROUP P[f]–PG[g]	
Recommendations	
Neutrino astronomy	10
Gravitational radiation experiments	10
Testing general relativity (ground-based)	10
TOTAL	30
GROUP P[f]–PS[h]	
Recommendations	
Gamma-ray detectors	15
Testing general relativity (space-based)[i]	30
TOTAL	45
GROUP L[j]	
Recommendations	
Scheduling of large optical and radio telescopes	–
Monitoring variable radio sources	–
Very-high-resolution studies	–

[a] Group A, of primary concern to Astronomy Survey Committee.

[b] AG, astronomy, ground-based.

[c] AS, astronomy, space-based.

[d] This refers to a series of smaller telescopes leading toward the Large Space Telescope. See *Astronomy and Astrophysics for the 1970's*, Volume 1, The Report of the Astronomy Survey Committee (National Academy of Sciences, Washington, D.C., 1971) for further discussion of this program.

[e] For x and gamma rays as well as for infrared.

[f] Group P, of primary concern to the Physics Survey Committee.

[g] PG, physics, ground-based.

[h] PS, physics, space-based.

[i] This is primarily for orbiting gyros to test inertial drag. The Panel believes that this program should receive careful study if the cost proves to be substantially higher than that given.

[j] Group L, low-cost recommendations. No cost figures assigned.

TABLE VIII-6 Estimated Costs of Recommended Programs (Ten Years)

Recommendation Group	Estimated Total Cost ($Millions)	Estimated Fraction, Astrophysics and Relativity	Estimated Cost, Astrophysics and Relativity Component ($Millions)
AG[a]	260	0.54	140.4
AS[b]	825	0.56	462.0
PG[c]	30	1.00	30.0
PS[d]	45	1.00	45.0
L[e]	–	–	–
TOTAL	1160		677.4

[a] Astronomy, ground-based.
[b] Astronomy, space-based.
[c] Physics, ground-based.
[d] Physics, space-based.
[e] Low-cost recommendations. No cost figures assigned.

plained above, astrophysics and relativity currently requires about $60 million per year, of which about $40 million is in space experimentation. The recommended budget would roughly double these figures over the next decade. It should be pointed out that the recommended budget is by no means the maximum that could be exploited at this time. A major new instrument, the Large Space Telescope, is *not* provided for in the recommended budget, although a program evolving toward it is recommended. It is believed that the Large Space Telescope might well cost a total of $1000 million or more. If we had assigned 30 percent of its cost to astrophysics and relativity, our recommended budget would have been larger by $30 million per year, or nearly 50 percent as a consequence of adding the Large Space Telescope. Many experts believe that such a telescope is feasible now. By not recommending an all-out program at this time, the Panel has taken a conservative approach, which may very well postpone the time when certain decisive cosmological measurements can be made by as much as a decade.

Let us consider the option of constant funding at the present level of $60 million per year. It is apparent from Table VIII-5 that expensive new facilities such as the large radio array, the diffraction-limited optical space telescope program, and the high-energy astronomy observatory could be undertaken even one at a time only by accepting deep cuts in the present program. As the present program is sound, this would not be a sensible procedure. We therefore assume that under a constant budget, there would be no major new instruments. One could, however, contemplate initiating some of the smaller programs, such as equipping major optical telescopes with electronic imaging, and expanding the present infrared program. Funds to do this would have to come out of present programs as well. It is hard to see how this could be done without closing down observatories that are presently productive, however. One area that would be hard hit under a level budget would be radio astronomy. No major new instrument would then be possible, in spite of the spectacular discoveries in this field and the fact that several proposals for new instruments have been in line for funding for five years or more. In summary, a constant budget would confine the subfield largely to observational techniques of limited power. It would not be possible to observe many faint phenomena, such as distant x-ray galaxies, infrared sources, or radio explosions, which are believed to be present and which have great significance for cosmology and for fundamental physics.

Decreasing support would of course prohibit any major new instruments and would also require that present major facilities be shut down. Although the situation is not desperate at present, there are hints that this in fact may become necessary in some areas, if the present paucity of funds continues. If this happens, one can predict that there would be a shift in the field toward theoretical work, and that innovations in instrumentation for new wavelength regions would all but cease. As we have tried to show in this report, this would reverse the trend that has made the field so exciting in the past two decades. Deprived of hard information on what the universe is really like, there is little hope that the theorist can arrive at the truth. Such questions as Is the universe finite or infinite? Are there new laws of physics operating in the quasars? How are particles accelerated to relativistic energy with very high efficiency in astronomical systems? Do black holes really exist? Does gravitational radiation exist, and what is its source? which there is some chance of answering in the next decade, given adequate support, will simply not be answered with a declining budget.

If, on the other hand, adequate support is available, it is believed that discoveries of profound significance will be made.

IX
Earth and Planetary Physics

PANEL MEMBERS

RICHARD M. GOODY, Harvard University, *Chairman*
DON L. ANDERSON, California Institute of Technology
W. I. AXFORD, University of California, San Diego
HENRY G. BOOKER, University of California, San Diego
A. G. W. CAMERON, Yeshiva University
JOSEPH W. CHAMBERLAIN, The Lunar Science Institute
WILLIAM GORDON, Rice University
ROBERT PHINNEY, Princeton University
WILLIAM RICHARDSON, Nova University

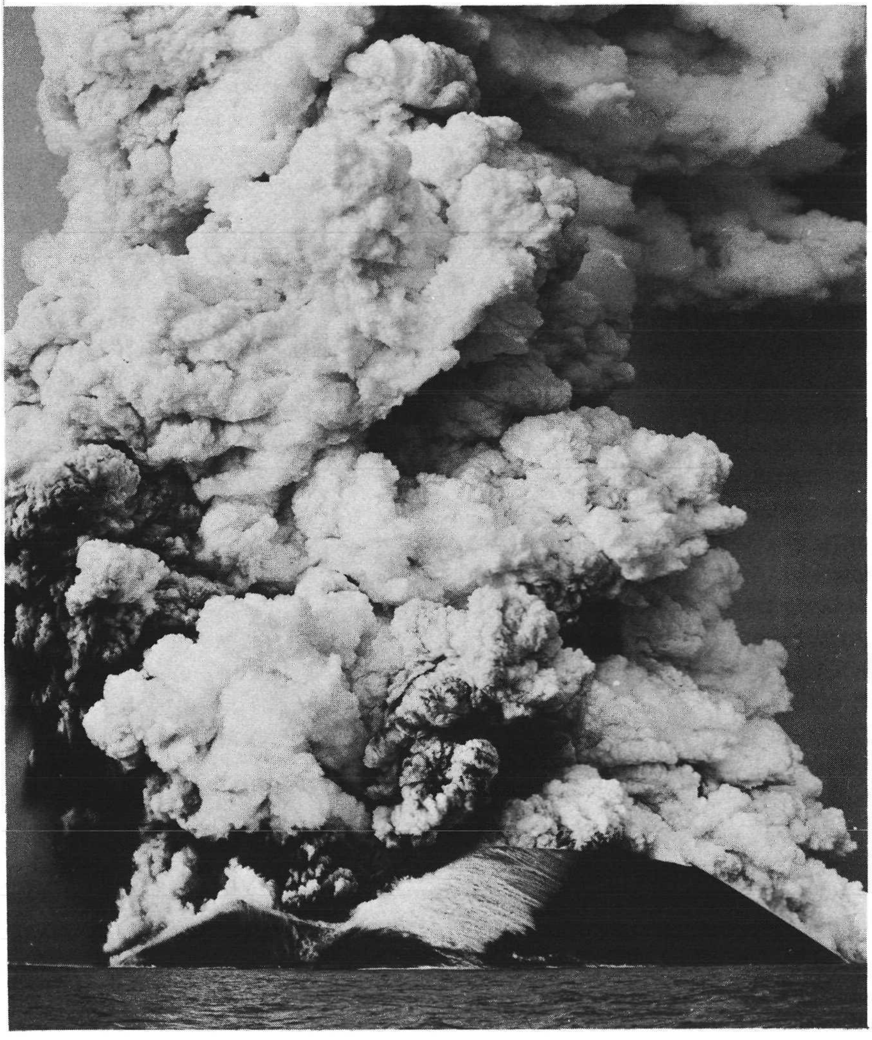
CONSULTANTS

ROBERT FLEAGLE, University of Washington
MARVIN GOLDBERGER, Princeton University
WALTER MUNK, Scripps College
FREDERIK ZACHARIASEN, California Institute of Technology

LIAISON REPRESENTATIVES*

DAVID B. BEARD, University of Kansas, Consultant to the Atomic Energy Commission
ALBERT E. BELON, National Science Foundation
NEIL BRICE, National Science Foundation
ROBERT F. FELLOWS, National Aeronautics and Space Administration
ALBERT OPP, National Aeronautics and Space Administration
ERWIN R. SCHMERLING, National Aeronautics and Space Administration

* In general, no agency was represented on the Committee or any of its panels by more than one person at a given time.



Air, water, earth, and fire. A submarine volcano erupts. [Photograph courtesy of Photoreporters, Inc.]

Preface

This report describes the use of physics in the earth and space sciences. Despite the wide differences in motivation among the different sciences involved, the research physicist sees them in a very similar light, and for disciplinary purposes they can be treated together. We shall stress that a physicist can work in the earth and space sciences only in full collaboration with many disciplines. Nevertheless, he is distinguishable by a style, his education, and his view of the phenomena of the physical world, which he interprets in terms of a few basic and simple laws. This loosely defined activity by physicists we shall identify as the physics subfield of earth and planetary physics.

This document consists of eight essays or commentaries addressed to the relationship between physics and the earth and space sciences. It was not possible for this Panel to follow the charge developed for the other panels of the Physics Survey Committee. As several of the following papers emphasize, earth and planetary physics is an important part, but only a part, of the earth and space sciences; the role of physics in these sciences cannot be examined as an independent factor. Nor can such matters as education, funding, and priorities be discussed adequately in the limited context of their involvement with physics.

In recent decades, the earth and space sciences have been subjected to exhaustive examination by the National Academy of Sciences, the federal government, and private organizations. The documentation is extensive, and many proposals have been developed. Merely to read all this material is a major task; this Panel did not see the need for, nor did it have the facilities to undertake, another review of the same character.

The Physics Survey, however, presents a novel opportunity. As Chapter 1 indicates, physics and the earth and space sciences are related, and close interaction between them is essential. Yet a separation has developed between physics and these disciplines, and there are many problems associated with this separation. The Panel felt that it would be valuable to identify and discuss these interface problems specifically. Therefore, the chapters that follow are addressed primarily to the physics profession. They describe the scope and character of the earth and space sciences and attempt to define their interface with physics in order to demonstrate the responsibility of physics as a scientific discipline in this multidisciplinary area.

Chapter 1 describes the general nature of the subfield. Two chapters then describe its research goals, first in terms of science (Chapter 2), and next in terms of applications to man's needs (Chapter 3). Chapter 4 is concerned with education, and the most important conclusions on the relationship between the earth and space sciences and conventional physics are developed here. The next three chapters deal with special features of the conduct of science in this subfield: Chapter 5 discusses national programs, facilities, and research centers; Chapter 6 is concerned with the special role of computers; and Chapter 7 describes international activities. Finally, Chapter 8 examines manpower, funding, and existing priority studies.

Because of the nature of this Panel's approach, we make few specific recommendations.

As this report goes to press we are uncomfortably aware that the year that has elapsed since the manuscript was completed has seen many changes, particularly in the NASA programs. The Office of Space Sciences and Applications has been split in two; Grand Tour has been replaced by a less expensive mission; the Space Shuttle has emerged as the post-Apollo program. In all agencies, budgets for fiscal years 1971 and 1972 are firm and to some extent funded, while fiscal year 1973 is before Congress. It was impossible to bring the manuscript up to date in all respects, but where changes were essential, we have tried to incorporate them, if only in footnotes.

We wish to acknowledge the invaluable help of Dale Teaney of the American Institute of Physics with statistics. We are also indebted to Frederik Zachariasen, David Rose, and Marvin Goldberger for opinions on environmental problems and the role of the physicist. We sought several expert opinions, which we were able to incorporate directly into the report: Jule Charney wrote the second part of Chapter 6; Hugh Odishaw prepared the whole of Chapter 7; James Baker provided material on oceanography for Chapter 3. Harvey Brooks read and criticized the manuscript, which, as a consequence, was greatly improved.

Contents

IX

EARTH AND PLANETARY PHYSICS

1 The Nature of the Subfield

The earth and space sciences are concerned with natural events occurring in the physical environment: in the interiors, atmospheres, and oceans of the earth and planets; in the outer layers of the sun; and in the interplanetary plasma. They deal with physical events; therefore, they properly belong to the physical sciences. Explanations are sought in terms of fundamental physical laws; physical scientific method, with some modifications to include observational techniques, is a powerful tool, and an education in physics, with some important additions, is a proper preparation for a research worker.

Before the twentieth century, there was little need to distinguish between physics and the earth and space sciences. Familiar figures in physics and mathematics invested time and effort in understanding atmospheric motions, atmospheric optical phenomena, ocean tides, the earth's magnetism, and the like. They were not the only people interested in the phenomena. The naturalist, with his ability to record and describe and his deep interest in natural events, was an important figure, but his activities were largely independent of those of the physical scientist.

In recent decades, physics, as conventionally defined in the United States, and the earth sciences have drawn apart—so far apart that often there is little or no communication. During this period, physics has tended to emphasize a core area of research into fundamental properties of matter. The earth sciences, on the other hand, have been faced with problems related to their applicability to man's material needs and, as a result, have developed educational and research programs directed to these ends.

The earth and space sciences have common features, described in subsequent chapters, but they also have great diversity. Only in recent years has their commonality been emphasized in the educational process and in federal activities. However, some insight is lost if the relationship between physics and the earth and space sciences is described solely in terms of their present status. A brief historical review of the different disciplines offers a sense of perspective and provides the opportunity to define the terms used in this report.

The science of the atmosphere has undergone as many changes as any of the earth sciences. Before the First World War, it was taken for granted that fundamental advances in the knowledge of atmospheric motions and the associated weather phenomena would eventually result from physical and mathematical research. During World War I, the need for weather information for aircraft operations became extremely urgent, and the telegraph and radio provided the means of collecting and disseminating the necessary information. Despite some brilliant but premature attempts, the success of weather prediction did not improve rapidly as a result of physical research; on the other hand, the Scandinavian school of synoptic analysts was able to demonstrate impressive advances using semiempirical frontal concepts.

At this juncture, a major distinction between meteorology and physics began to emerge. Separate departments and research centers for meteorology developed in universities, often with curricula weak in physics and mathematics and an emphasis on descriptive methods. The interests of the physicist in atmospheric phenomena began to be described as atmospheric physics, and in the 1950's upper-atmospheric physics was distinguished by the term aeronomy. For a brief period, a serious dichotomy between lower- and upper-atmospheric research began to develop, with the former emphasizing descriptive and synoptic methods and the latter, the methods of physics. Even at that time, however, the advent of electronic computers and advances in the knowledge of the physics of large-scale hydrodynamics, radiative transfer, clouds and precipitation, turbulent boundary layers, and the like showed that problems of the lower atmosphere could also be treated in terms of the fundamental laws of physics. Simultaneously, increasing observations of the upper atmosphere showed it to be little, if any, less complicated than the lower atmosphere and in need of meteorological experience and techniques.

As a result of these trends, the concept of the atmospheric sciences was introduced in the 1960's to describe the synthesis of all approaches to atmospheric investigations, the development of atmospheric sciences curricula in universities, with stronger emphasis on mathematics and physics, and the creation in the United States of federal programs and research centers to support all aspects of atmospheric research.

The term "oceanic" or "marine sciences" also conveys the notion that physical, descriptive, biological, and other disciplines must be combined to solve the problems of the oceans. The marine sciences are similar in many ways to the atmospheric sciences. The major differences are a shorter history of intensive funding; a large involvement by life scientists; and more freedom of approach, because of the less pressing demand for immediate applications. The terms "oceanography" and "physical oceanography" express a parallelism different from that between meteorology and atmospheric physics, but the reduced role of applications has made the difference less significant than in the atmospheric sciences.

From the disciplinary point of view, physical oceanography and meteorology have been very close. Ideas can readily be exchanged between the two disciplines, particularly in fluid dynamics. Oceanography does not have the observational basis of meteorology and has had to draw many of its ideas from atmospheric research. As a consequence, the two subjects are often handled in the same university department, and some research workers are active in both fields.

Geophysics, or physics of the solid earth, has evolved through three distinct traditions. The early work on the mechanics of the earth spawned continuing studies by a few independent physicists, affiliated variously with physics departments or research institutes. In the same way, a few individuals in geology departments rejected the descriptive approach and sought a physical basis for their subject. The earliest formal groups in the United States were the Department of Geophysics at Saint Louis University, the Carnegie Institution's Department of Terrestrial Magnetism, and the Caltech Seismological Laboratory, all founded in the early 1900's. At that time, geophysical methods were successfully adopted for oil and mineral exploration, well in advance of full theoretical understanding. From these separate origins arose the many and diverse modern research groups. Those located in universities may be affiliated with geology, physics, or mineral engineering departments. Some of the most successful are amalgamated in autonomous research institutions, in which the different traditions exist together under one roof.

Geophysics has undergone less rapid change than other earth sciences because of the relatively moderate growth curve of research and the lack of support for large, coordinated projects. Although the International Geophysical Year 1957–1958 provided an important stimulus, the first real injection of large-scale government support in solid-earth studies came as recently as 1960 in the form of the VELA program for detection of underground nuclear explosions. The research capabilities created as a result of this support are now being turned toward important new problems related to earthquake prediction and control.

A major advance in solid-earth studies came about in 1966 as a result of a wide-reaching collaboration between ocean-going and land-based workers. The study of geomagnetic polarity reversals preserved in lavas, the determination of earthquake source motions, and the age-dating of continental basement rocks have shown conclusively that the earth's outer shell is extremely mobile on geologic time scales. Earth scientists are on the threshold of a completely new era of solid-earth geophysics, in which major directions in research appear to change daily.

A new emphasis in geology departments is developing as studies of planetary surfaces become feasible and as more lunar samples are returned, but this influence is by no means preponderant. However, the relationship of the solid-earth geophysicist to geology and physics departments is much the same as that of his colleagues in the oceanic and atmospheric sciences; he seeks the motivation, the data, and the problem definition from one and the necessary scientific techniques from the other. Although exclusive concentration in either direction is clearly unsatisfactory, no agreement exists as to the optimal synthesis.

Terminology in solid-earth studies can lead to confusion. In this report, we shall use the term "geological sciences" to include geophysics and all classical aspects of geology. The term "earth sciences" is sometimes used as a synonym for geological sciences, but we shall use it to denote the entire group of sciences concerned with the atmosphere, the oceans, and the earth. That part of the earth sciences that involves the use of physics we shall call "earth physics." The term "geophysics" is sometimes used in this broad context, but we shall avoid such usage.

A further problem in nomenclature has been introduced by federal government use of the term "environmental sciences" to describe almost the same group of activities that we include here under earth sciences. This nomenclature is misleading because, although the earth sciences are concerned with the environment, the general public accepts a much wider connotation, including ecology. However, the term necessarily occurs in any discussion of federal budgets.

This report also is concerned with certain branches of the space sciences. During the late 1940's and the 1950's, when the only vehicles were rockets, aeronomy was the frontier field of space science. The rapid accumulation of detailed data using rockets and later satellites has been largely responsible for the convergence with ideas previously associated with lower-atmosphere studies.

The advent of the satellite in the 1960's led to the discovery of the earth's radiation belts and the interplanetary plasma. These studies found a secure home in physics departments because of their association with cosmic ray research. They are largely empirical, but of a novel character, and are almost entirely dependent on the availability of space probes. A change in the policy of the National Aeronautics and Space Administration (NASA) could virtually eliminate this field of study. Therefore, it is an exceptional field and difficult to relate to the others with which this report is concerned. According to NASA terminology, aeronomy and the interplanetary plasma together constitute the subfield of space physics.

Before 1960, studies of the planets were considered a branch of astronomy. However, astronomers and departments of astronomy often showed an equivocal attitude toward this field, sensing, perhaps, an essential difference of approach. Astrophysics is largely concerned with matter at extremely high temperatures and pressures under conditions in which the laws of physics established in the laboratory may fail. Planetary studies, however, are concerned with conditions close to those of the earth, and the problems involved are likely to have parallels in terrestrial studies. When observing techniques for planets were the same as those used for astronomy and when only crude data were available, this parallelism was not exploited, and few earth scientists were involved in planetary studies.

The lunar landings and unmanned planetary probes have now begun to supply data of the certainty and detail needed to apply the methodology of the earth sciences, and most of the work on planets (planetology, planetary meteorology, planetary aeronomy) is now performed by geophysicists, meteorologists, and aeronomers. Earth physics and planetary physics have become part of the same discipline; departments and research centers of earth and planetary physics are appearing in the universities; and the relationship between planetary research and astronomy becomes increasingly distant.

The gap that exists between physics and the earth and space sciences operates to the disadvantage of all concerned. But any change in the present circumstances depends on an understanding of factors that are difficult to identify. These factors include differences in fundamental approach, differences between observational and experimental methods, differences in the emphasis placed on practical goals, and the interdisciplinary nature of research in the earth and space sciences, in which physics is a major, but not the only, contributor.

Some of the essential differences between the earth sciences and physics or astronomy were identified by Aristotle.[1] He contrasted the earth sciences with investigations of "the first causes of nature" (physics) or "the stars ordered in the heavens" (astronomy) and stated that they are concerned with "events that are natural, though their order is less perfect than that of the first of the elements of bodies"; he proposed to seek explanations following the phenomenological methods used for "animals and plants."

With the advantage of hindsight, we know that Aristotle's distinctions among the three disciplines are not as clear as his definitions suggest. The remarkable advances of physical science leave no doubt that events taking place in the solar system can be explained in terms of physical laws established in the laboratory. The problem is one of unraveling a complex of interacting phenomena on scales of space and time not usually encountered in the laboratory and generally without the help of the controlled experiment.

The earth and space sciences are observational sciences—they obtain their data by direct observation of natural systems rather than by experimentation. There has been a tendency to define physics strictly as an experimental science and thus to distinguish it from these sciences and from astronomy. The technique of gaining understanding by observing a system with many parameters evolving in space and time must be learned, just as the experimental method must be learned. There are a few problems in the earth and space sciences for which laboratory experiments are important (the work of Hide and Fultz on rotating fluids is a good example); in attacking these problems, the physicist is on more familiar ground. But even in these rare cases, a great deal of experience is required to relate the experiment and the natural phenomenon.

Not only are observations the basis of research in earth and planetary physics, but the observations are numerous, sophisticated, and difficult to view with a sense of proportion. A certain naïveté can often be productive, but, in the long run, a physicist, to be effective, must also gain some of the experience and judgment of a meteorologist, a geologist, or an oceanographer, and he must learn to communicate with a number of other disciplines.

Criteria of success also differ from those of conventional physics. In earth and planetary physics, the success of a theory is not judged in terms of new insights into fundamental laws; rather, such physicists are concerned with the number of observed phenomena that can be related in terms of a hypothesis based on the known laws of physics. In this respect, earth and planetary physics has something in common with astrophysics, although recent astronomical discoveries require completely new ideas about the nature of matter at high temperatures and high pressures.

Although the earth and space sciences are properly described as physical sciences, their procedures often are influenced by considerations not strongly sensed in the more conventional areas of physics. The earth sciences are largely, but not exclusively, concerned with questions immediately related to needs of society, for example, weather forecasting. To answer these questions, deductive scientific method is used to the greatest extent possible, but the problems involve so many poorly understood interlocking phenomena that immediate answers must be based partly on experience and judgment. This situation is parallel in many respects to that in medicine, and the consequence in both is the emergence of a large body of practitioners whose activities the general public supports.

The relationship between practitioner and scientist is not hierarchical: it is not the case that one performs the fundamental research and the other applies it. Information flows in both directions. The practitioner, of course, must be able to understand scientific advances and to make as much use of them as possible; but, on the other hand, the data collected for empirical analyses are an essential source for the scientist denied the use of the controlled experiment. The present knowledge of atmospheric circulations could not have been reached without the data collected for weather forecasts; physical oceanography has long depended on data assembled to aid ships' navigators; geophysics would have no point of departure without the data collected by seismologists, geologists, and others.

This emphasis on applications and on the procedures needed to achieve results on a short time scale is perhaps the primary reason for the divergence between physics and the earth sciences. The questions with which this Panel is concerned are, therefore, in part characteristic of the dichotomy between pure and applied physics research as it has developed in the United States, and the problems arising in regard to manpower and educational matters are not limited to the earth and space sciences. It should be pointed out, however, that the gulf between the core area of physics and the earth and space sciences is made unnecessarily wide by the way in which physics is defined.

First, both disciplines, for example, space physics, and individual geophysicists, meteorologists, and oceanographers are satisfactorily accommodated in physics departments and research centers in the United States. Moreover, the division between physics and the earth sciences is more strictly observed in the United States than in Europe or the Soviet Union. In the United Kingdom, for example, there are relatively few departments of meteorology, oceanography, or space sciences, and physicists working in these disciplines in the universities are usually found in physics departments. This situation may be responsible in part for what sometimes appears to be a higher quality of research in relation to the total number of people engaged. In Germany and Scandinavia, the relationship between physics and the earth sciences differs from that in the United Kingdom, but they are still much closer than in the United States. The Soviet Union pays less attention than most countries to the distinction between pure and applied research; atmospheric physics, for example, is a major and continuing activity of the physics department at the University of Leningrad.

An important characteristic of the earth and space sciences is the extent to which they are concerned with phenomena of unusual aesthetic appeal and with humanistic values related to man's place in nature. In his history of the study of the rainbow, Boyer remarks that "meteorology (in the mediaeval sense) formed an important part of mediaeval scholasticism."[2] His case is well documented by a host of references on sky halos alone. One of these, the rainbow, commanded the attention of many of the greatest minds in science until the essence of the problem was solved by Descartes, Newton, and Airy. This fascination with atmospheric optics related as much to aesthetic as to scholarly values.

The same is true of many areas of investigation in the earth sciences. The visual impact of the aurora polaris stimulated much work; it is difficult to be unmoved by satellite pictures of the motions of clouds in a mature hurricane; the planets have always been considered among the most wondrous of natural objects; and men look with more appreciation and enjoyment on the surrounding landscape when they understand something of the geological forces that formed it. (See Figures IX.1, IX.2, and IX.3.) Newton may have been accused of destroying the poetry of the rainbow by reducing it to prismatic colors; but, in terms of contemporary thinking, the arts would seem only to have gained by physical understanding of light, color, sound, and events in man's environment.

The earth and space sciences have to do with natural events from a human perspective. Their relevance to man's material and economic affairs has been emphasized, but the relationship is more intimate than is obvious at first, for the knowledge gained from these studies can be used to plan and proportion human affairs so that they are compatible with the natural world. The recent demands for pollution control for purposes of health as well as the aesthetic quality of the surroundings bring politics, law, sociology, and economics into interaction with medicine, hydrology, meteorology, and oceanography in an attempt to achieve a desirable change in human affairs. The people who are involved in the direction of human affairs must become aware of the methodology by which these large geophysical problems are handled if they are ever to bring the human component successfully into their studies. It can be argued that the observational sciences provide a route by which the exact ideas of the physical sciences can be made useful in the complex area of human behavior. For example, L. F. Richardson made important contributions to the theory of sociology after he decided to abandon his pioneering work on the numerical prediction of weather, a decision made because he could obtain research support only from a military agency.

FIGURE IX.1 Apollo 17 view of earth. (Photograph courtesy of NASA.)

FIGURE IX.2 The Barnard Glacier, Alaska. (Photograph courtsey of Bradford Washburn.)

FIGURE IX.3 Deep-ocean swell. (Photograph courtesy of Jan Hahn, Woods Hole, Massachusetts.)

Finally, we come to man's more remote strivings to understand his origin and that of his world and universe. The idea that such understanding rests exclusively on belief rather than knowledge has largely disappeared from modern societies, and people in all walks of life have an unlimited appetite for new facts about these matters. Not without reason did the Space Science Board of the National Academy of Sciences–National Research Council emphasize the search for the origin of life and the origin of the solar system as two of the major justifications for NASA's lunar and planetary program.

So far, we have stressed the differences between physics and the earth and space sciences rather than their considerable common ground. The relationship between these two branches of the physical sciences should be stronger than it is: The earth and space sciences need a greater influx of trained physicists and new ideas from physics, and physics, as traditionally defined, can also benefit from the association, for the earth and space sciences have certain aesthetic and humanistic values to offer. The young physicist who is not acquainted with the earth and space sciences could well miss an opportunity to work in physical research outside the laboratory, to contribute work and knowledge that are closely related to the needs of society, to observe natural phenomena that are often of great aesthetic appeal, and to attack challenging problems of unusual difficulty and complexity. These considerations will not appeal to all physicists, but they do appeal to some and they could appeal to more if a broad view of physics as a discipline were the objective of our educational system.

2 Scientific Goals

A common thread runs through many of the prime scientific goals in the earth and space sciences. The advances in instrumentation, theory, and computing power during the scientific buildup of the last decade have brought earth and space scientists to the point at which complete, positive knowledge of many aspects of the environment can be expected. To attain this kind of understanding requires a variety of well-planned observational programs. In addition, meaningful interpretation of the data thus obtained requires that people of the highest caliber be brought into these sciences to attack such problems as the origin of the earth's magnetic field, the nature of turbulence in stratified systems, and the equation of state of silicates at high pressures.

In this chapter, we describe some of the problems of greatest current scientific excitement and promise. Because of the nature of this subfield, many of these problems appear to be slanted toward applications. However, this discussion presents the scientist's point of view. Other chapters deal with the way that the profession can best respond to current environmental demands.

The discovery in the 1960's of sea-floor spreading (see Figure IX.4) and large-scale motions of the continents was the result of a broad range of geophysical and geological studies in the ocean basins and on land. Random polarity reversals of the earth's dipole magnetic field at intervals of 200,000 to 1,000,000 years are permanently recorded as magnetization in newly formed lavas and can be detected as short-wavelength elements of the field by a shipborne or airborne magnetometer. The geomagnetic polarity time scale for the last five million years was determined by direct study of radiometrically dated lava samples from several locations. Magnetometer surveys revealed a striking pattern of striped anomalies in the ocean basins; these anomalies (up to 50 km across) are symmetric with respect to the rift valley at the axis of the midocean ridge and correlate in detail with the polarity reversal time scale. The scale factor in this correlation is just the rate at which new crust is being formed at the rift, or equivalently, the rate at which two segments of the ocean floor are moving apart.

With observed spreading rates of 0.5 to 10 cm/yr, it is evident that the world's ocean basins are newly recreated on the order of every 250 million years. Direct evidence of this comes from the JOIDES* deep-drilling program, which uses the age of the deepest marine microfossils in a core to date the crust. The oldest oceanic crust (250 million years) is found east of Japan. Destruction of the crust occurs by sinking of the leading edge of a plate at the zone of convergence marked by the oceanic trenches. The effects of this collision and sinking of crustal plates appear to be responsible for the complete panoply of geological and geophysical phenomena in these regions: mountain building, earthquakes, volcanism. The lower-density continental rocks are elevated by hydrostatic balance and remain at the surface through this process, acquiring a permanent record of old plate motions through mountain belts, which become patched to the edge of the continent by collision with the fast-moving, ephemeral oceanic plates. Where two continents collide, the effects are spectacular; the collision of India with Asia is responsible for the doubling of crustal thickness in the Tibetan plateau as well as the great uplift of the Himalayas.

The discovery of sea-floor spreading has provided geophysicists and geologists with a great unifying principle. The greatest priority for the next decade or more is to collect the crucial data that are necessary to develop for all the phenomena a detailed model of the mechanisms that drive and

*JOIDES (Joint Oceanographic Institutions Deep Earth Sampling) is a program to core the oceanic crust, principally the 0.5 to 2 km of sediments that lie atop the basaltic lavas. The JOIDES program has operated for 3 years, obtaining to date 200 holes in the Atlantic and Pacific Oceans, the Mediterranean and Caribbean Seas, and the Gulf of Mexico. Participation of foreign scientists, including some from the Soviet Union, has been an important aspect of this program.

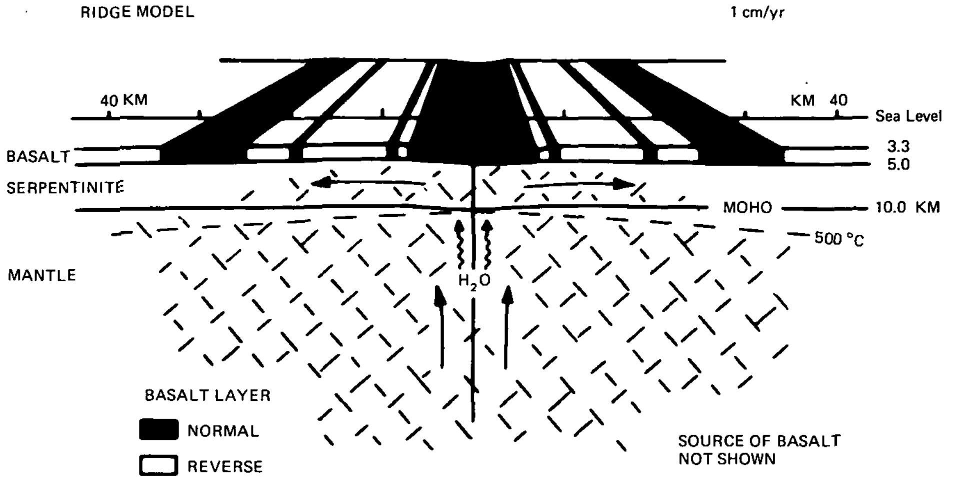


FIGURE IX.4. Sea-floor spreading.

limit the great terrestrial heat engine. Measurement of earthquake signals over 15 orders of magnitude of energy gives detailed information on the mode of slip and strain release along the boundaries between crustal plates. At sea, magnetic surveys, acoustic and seismic profiling of crustal structure, and heat flow and gravity measurements will continue to be important. On land, seismic holography (in some practical version) is needed to provide accurate three-dimensional reconstructions of the structure of the crust and upper mantle.

Because of their dependence on field measurements in diverse and inconvenient places, students of the solid earth are faced with the high experimental costs familiar to the physicist but are usually unable to achieve the economy that the physics or space community experiences in a large centralized facility. The JOIDES and WWSSN* programs are the only excep-

* WWSSN (World-Wide Standardized Seismographic Network) was established by the Environmental Science Services Administration of the National Oceanic and Atmospheric Administration, Department of Commerce, in support of the VELA nuclear test detection program and of general scientific research in 1961. It consists of 115 standard stations (long- and short-period three-component seismometers) and a central data facility for the microfilming and dissemination of records. In recent years, the funding of this network and its upgrading have become precarious.

tions to this situation. A major need for this decade is to develop unifying programs that move ahead in a coordinated way but retain the traditional institutional independence. The Geodynamics Project is an international effort sponsored by the International Unions of Geodesy and Geophysics and of Geological Sciences that promises to provide leadership and communication; a corresponding program of funding from the U.S. Government, however, is still under discussion.

The scientific achievements of the space program have brought planetology into being. Continuing studies of the earth's interior composition and evolution appear in a new perspective as data on the structure and composition of the moon and planets become available. Comparative study of the planets leads to inferences about the conditions of their formation. Compositional fractionation can be seen in the varying mean densities of the inner planets, which reflect principally the varying proportions of iron and its oxidation state. A puzzling fractionation of water seems to be indicated by the very small proportion of water to carbon dioxide in the atmospheres of Venus and Mars compared with earth. Differences in thermal evolution of the planets reflect both initial temperatures and the subsequent distribution of heat-producing radioactive nucleides within each planet.

The moon, from which new data can be cited, provides an example. At present, the moon has about ten orders of magnitude less seismic energy release per gram than the earth. Localized mass anomalies near the surface of the moon, inferred from the perturbations of satellite orbits, are of such a magnitude that the outer shell of the moon must be relatively cool (less than, say, 400°C) to a depth of about 400 km (see Figure IX.5). It is known from laboratory studies of rock deformation at high temperatures that these nonhydrostatic loads must otherwise subside in a few million years or less. Analysis of the inductive response of the moon to magnetic disturbances in the solar wind, obtained by comparing magnetometer signals from an Apollo surface instrument and an orbital instrument, leads to the inference of a cold (low-conductivity) outer shell, surrounding a warm (high-conductivity) core. Whether major portions of the core are partially molten is still debatable; it is especially interesting in light of evidence from the Apollo samples that all but possibly a very minor proportion of lunar volcanism occurred prior to 3 billion years ago. The picture that emerges is that of a history uniquely different from that of the earth. While the earth has been in a state of thermally driven convection and crustal rejuvenation for at least 3.5 billion years, the moon, after an early phase of thermal activity, has lain geologically dormant. The distinction is seen in the thickness of their respective rigid shells—for the earth, about 60 km, or $0.01R$; for the moon, about 400 km, or $0.25R$.

The frontiers in planetology lie equally in continuing planetary explora-

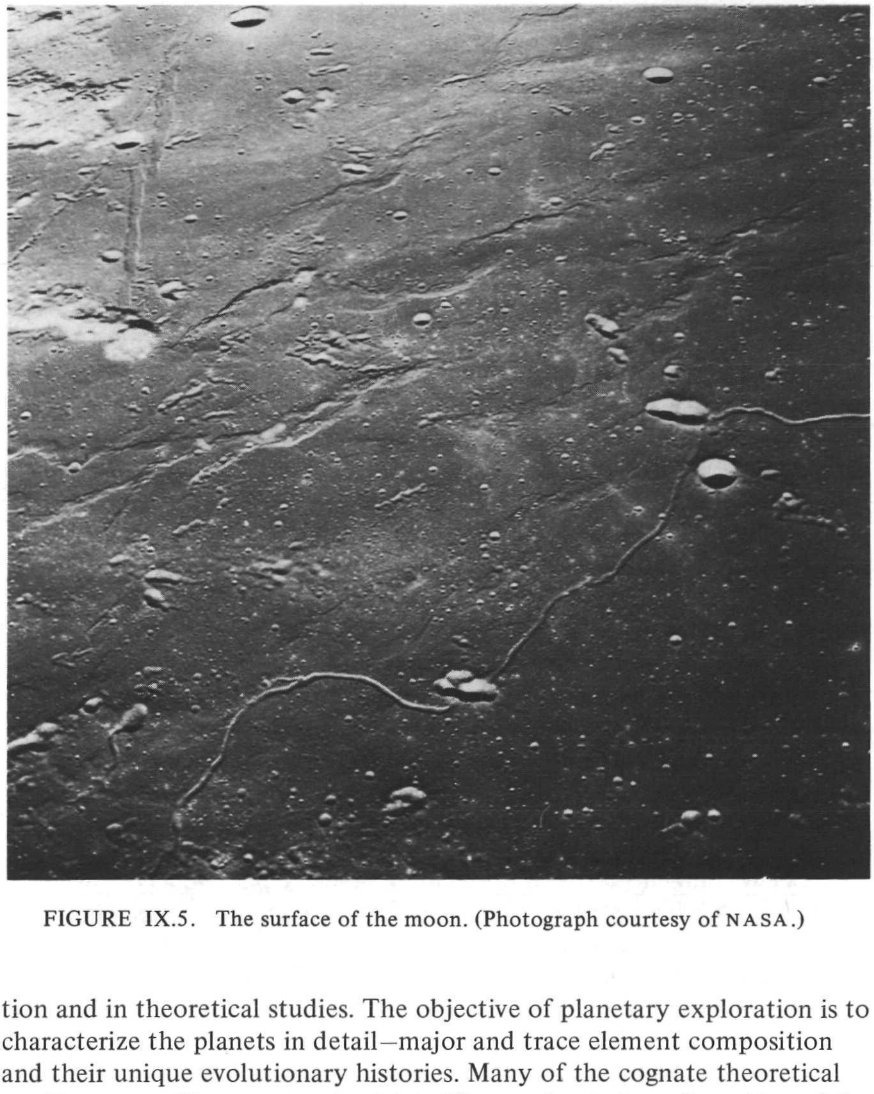


FIGURE IX.5. The surface of the moon. (Photograph courtesy of NASA.)

tion and in theoretical studies. The objective of planetary exploration is to characterize the planets in detail—major and trace element composition and their unique evolutionary histories. Many of the cognate theoretical problems are of interest to physicists. The mechanical configuration of the solar system can be described for short times by the laws of celestial mechanics, by equations for the periodic and secular perturbation of the Keplerian orbits. For long intervals, these equations are ill-posed, due to the existence of strong interactions at rare times, such as capture or ejection by Jupiter. In this setting, the dissipative tidal interactions between bodies seem to be a determining factor in the secular evolution of many objects. Tidal forces have been important in both the earth–moon system and the

Venus–sun–earth system. It still is not possible to handle these equations with any degree of realism. In addition, the vital dissipation term in the equations depends in detail on poorly known mechanisms in the interior of the planet.

Magnetic fields are manifested both as present-day fields (earth, Jupiter) and through remanent magnetization of rocks (earth, moon), which records past magnetic or geological events. Current theoretical work is concerned with hydromagnetic mechanisms for field generation. The equations for the earth's core, for example, are strongly nonlinear and have low symmetry. Numerical study of these equations requires an order of magnitude more computation than the numerical weather prediction problem and is still largely untouched. The equations for the Jovian field are formally similar, but the setting is different, with circulations in the outer regions of the planet believed to be responsible. The Jovian problem will be attacked by spacecraft, such as the Pioneer F and G and the outer planets missions, which will monitor the magnetic fields close to the planet, obtaining high-resolution information on the high-order (wavenumber) components, and which will use imaging to improve the hazy picture of the atmospheric circulation. The earth's core is not fully explored by indirect methods, either. High-resolution analysis of seismic waves scattered or diffracted by the core in certain directions has shown nonradial variations of the sort that might be related to the current hydrodynamics of the core. For either planet, the problem can be solved only by an astute combination of theoretical analysis and inference from limited data.

Magnetic fields also have become particularly important in the interplanetary plasma and the earth's magnetosphere, in which they are a dominating factor in determining the dynamic state of the plasma. The study of the earth's environment in space, in the interplanetary plasma, has grown more than any other field of research as a result of the space program. The normal interplanetary medium is a spiral stream of ionized matter thrown out of the solar corona, characterized by particle densities of 5–100 cm^{-3}, velocities of about 500 km/sec, and frozen-in magnetic fields of about 5×10^{-5} G. (See Figure IX.6.) Direct sampling of the solar wind plasma has begun to give direct evidence of the solar composition. Spacecraft with particle spectrum analyzers have improved from an initial sensitivity to protons and alpha particles to the capability of detecting and counting heavier nuclei. The solar wind-foil experiment, deployed on the moon during Apollo EVA's, has detected up to $Z = 18$. The surface layers of moon rocks have captured a remarkably detailed record of solar wind and solar fluxes, through direct capture of rare gases and the characteristics of particle tracks in the crystals.

The nature and behavior of the solar wind will remain a major research

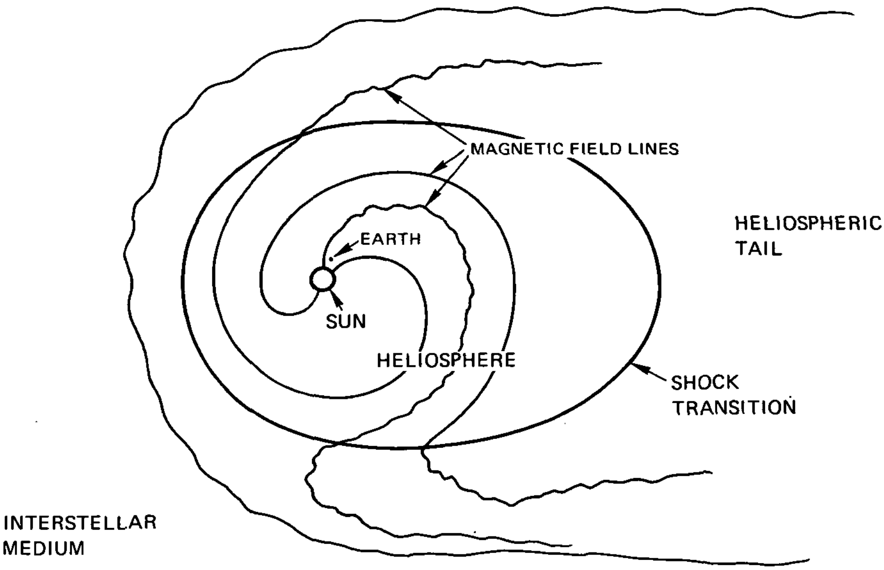


FIGURE IX.6. The concept of the heliosphere.

objective, both for its own sake and as a plasma laboratory of enormous scale. However, as space probes for the outer planets are developed, interest is shifting to the feasibilities of the direct observation of galactic matter. Galactic cosmic rays are observed on earth but in a modified condition for energy less than 1 GeV, because of the interaction inside the heliosphere with solar wind magnetic fields. The transition from solar to stellar material should take place somewhere between Jupiter and Pluto, that is, at distances accessible to outer planets spacecraft. From the composition and energy spectrum of unmodified cosmic rays, scientists should begin to understand their origin, age, and mode of propagation and should have direct information pertaining to their stellar sources.

The interaction of the interplanetary medium with the earth is just one of many very complicated, interrelated problems concerning earth's upper atmosphere. This region, which is regarded as extending from about 100 km out to 14 earth radii or more, is characterized by high temperatures, low particle densities, ionization, and rapid fluid motions. Above 300 km, the low particle densities, long mean free paths, and the strong terrestrial magnetic field characterize the magnetosphere, with its energetic particle belts. This entire region is of great scientific interest because of the rich range of phenomena that are found there.

The immediate scientific priority in the study of the outer atmosphere and magnetosphere lies in the adequate quantitative characterization of the full range of phenomena present—adequate to give an understanding of the physical processes that are important. Only then can any predictive modeling be contemplated. For example, the conservative dynamics of particles in the radiation belts are easily explained; the processes by which the particles are injected and accelerated and by which they decay involve fairly complicated interactions with the atmosphere and the solar plasma and are critical in determining the energy spectra and time constants.

This work was originally motivated by the most spectacular of all atmospheric phenomena—the polar aurora (see Figure IX.7). Although scientists understand clearly many aspects of particle acceleration and excitation of optical phenomena by high-energy particles, they still have no understanding of the essential link, the precipitation mechanism. The mystery was deepened in recent years by the discovery of the infrared oxygen aurora, with such high intensity that it greatly exceeds all known and proposed sources of energy in the magnetosphere.

Given the cost of direct measurement and the finite resources in the space program, the intelligent, economical design and selection of experimental payloads is of paramount importance. Ground-based methods of observation will be improved to exploit the possibilities of monitoring the outer atmosphere by using more subtle, indirect methods. A proposed national facility for incoherent radar backscatter is an example of this kind of effort. Radar signals backscattered by thermal fluctuations in the electron density can be used to infer a wide variety of properties of the plasmasphere and ionosphere: electron density, electron temperature, ion temperature, ion composition, mean plasma drift velocity, ion–neutron collision frequency, and the direction of the earth's magentic field. Electron density influences the scattered signals at all altitudes and can be determined in a number of separate ways. The other parameters manifest themselves by altering the shape of the power spectrum of the scattered signal. Since these measurements can be made at frequent intervals by incoherent-scatter radars, this technique becomes an almost ideal probe for the upper atmosphere, providing good altitude coverage, time coverage, and resolution at a particular location. These studies provide such a spectrum of data to students of the upper atmosphere that incoherent scatter facilities are truly general purpose. For example, drift velocity determinations provide important inputs to the continuity equation governing the ionosphere. Measurements along the magnetic field of density, temperature, and velocities at high altitudes provide estimates of particle and heat fluxes into and from the protonosphere and are used to establish boundary conditions for the study of the *F*-region ionization.

Radio propagation studies are, of course, the classic indirect probe of the

FIGURE IX.7. The aurora. (Photograph courtsey of Victor P. Hessler.)

ionosphere. Passive methods, using naturally occurring low-frequency signals, still remain an underexploited approach to probing the dynamics and state of the outer atmosphere. Whistlers, transient electromagnetic signals in the acoustic frequency band, and micropulsations, random magnetic signals in the 10^{-3}–10^{-1} band, have yet to be fully understood and used. The micropulsations involve both hydromagnetic and low-frequency electromagnetic modes of the earth–ionosphere–magnetosphere system. Their spectra contain information on both the characteristics of the system and the particular dynamics that are exciting the system.

The pressing need, as in other related types of studies, is for appropriate and modern observing networks in close combination with a theoretical effort aimed at rationalizing the data with models of the system.

The *D*-, *E*-, and *F*-layers of the ionosphere used to be the prime focus of attention in upper-atmosphere physics. They remain objects of interest, but the main features and basic mechanisms seem to be understood in principle, except for the dynamical transports by winds and turbulence, a subject that this field has in common with lower-atmosphere studies. An acceptable quantitative theory of the extent of the mixed region of the ionosphere—the turbosphere—would be one of the most important advances that could be made in the field.

The focus of attention in ionospheric theory has recently been on the comparison of earth, Mars, Venus, and the outer planets. The application of ideas developed for earth to other planets is a crucial test of their validity. The predicted *F*2-layer on Mars and Venus was not there when direct observations were made, leading to the problem of why their atmospheres have virtually no atomic oxygen at high altitudes. According to terrestrial theory, this should not be so, and scientists are still seeking new ideas about the photochemistry of carbon dioxide to account for it.

The planets also offer a unique proving ground for ideas about interactions of the solar wind with planetary ionospheres. Mars and Venus have no magnetic fields; consequently, scientists are trying to discover how the ionized plasma penetrates and what influence it has on the ionized layers gravitationally bound to the planet.

During the two and a half decades since the end of the Second World War, steady advances have been made toward one central goal of meteorological research: understanding and predicting weather phenomena on the basis of physical laws. What might appear at first sight to be no more than an integration of Newton's laws as applied to a fluid in fact involves physical and mathematical techniques that were not available when Richardson first worked on the problem in the 1920's. Many important components of the problem still are not understood; however, there is no longer any doubt that all relevant physics relating to scales larger than the computational grid size (a few hundred kilometers) is understood in principle and can be incorporated in practice, at least for middle latitudes. But half of the atmosphere lies in tropical regions, and phenomena of less than grid size are as important as any aspect of terrestrial weather.

Examples of important phenomena of relatively small scale are the front and the tornado. Fronts obviously are related to major circulation systems in the lower atmosphere, and no theory that does not predict their occurrence and intensity can be considered complete. However, at the present time, there is no generally accepted theory of frontogenesis, and relatively small effort is currently invested in the field. Much the same can be said about the tornado, but in this case even an adequate observational definition is lacking and the technology required is not in sight. Even the related problem on a smaller and more tractable scale—the dust devil—is neither observed nor explained in a satisfactory manner, and such phenomena may be of more than academic interest. On rapidly rotating planets such as earth, and especially Mars, small, rotating convective elements may, at times, be the principal means of heat and momentum transport across the planetary boundary layer. In other words, it is quite possible that much of the work so far performed on boundary layers is simply inapplicable to major meteorological problems.

Many motions of comparatively small scale are arbitrarily designated as turbulence and treated by means of parametric measurements that have their origin in laboratory and engineering studies of homogeneous and isotropic situations. The existence of these studies and of Taylor's aesthetically satisfying statistical formulation may well have held the subject back in recent years. Attempts to develop atmospheric theories along parallel lines have been unproductive except at rare observational sites where suitable conditions occur. It now begins to appear that the irregular character of turbulent motions is not their most important feature and that greater attention should have been given to the nature of more organized dynamical elements. The dust devil is one such phenomenon, and, on a larger scale, success has been achieved in a number of cases by assuming the turbulence to have the transfer properties of the most rapidly growing or most efficient instability that the system can possess. These ideas have been applied to the general circulation as a whole, to laboratory convection experiments, and to the Eckman boundary layer in the atmosphere. They represent a complete break with the past and are stimulating new ideas with implications for all of the earth sciences, astronomy, and engineering.

Attempts to extend forecasts up to 14 days have drawn attention to the extraordinary neglect of tropical meteorology. Almost all known dynamical theory applies to midlatitudes, where rotational constraints are very important. Strong Coriolis forces are responsible for many of the characteristics of midlatitude weather, and their reduced magnitude in tropical regions is the main cause of the differences in meteorological characteristics. The transfer of heat and momentum by cumulus convection, the importance of direct overturning convective motions (Hadley cells), the nature of the intertropical convergence zone, and the stability of tropical zonal flows—these and many other matters are scarcely understood, and their comprehension on the same level as midlatitude phenomena is a major scientific goal for the 1970's and 1980's.

In the field of tropical meteorology, one problem stands out above all others—the hurricane or typhoon (see Figure IX.8). Scientists are still in the process of observational definition, and no theory is generally accepted. At the same time, attempts are being made to influence the path

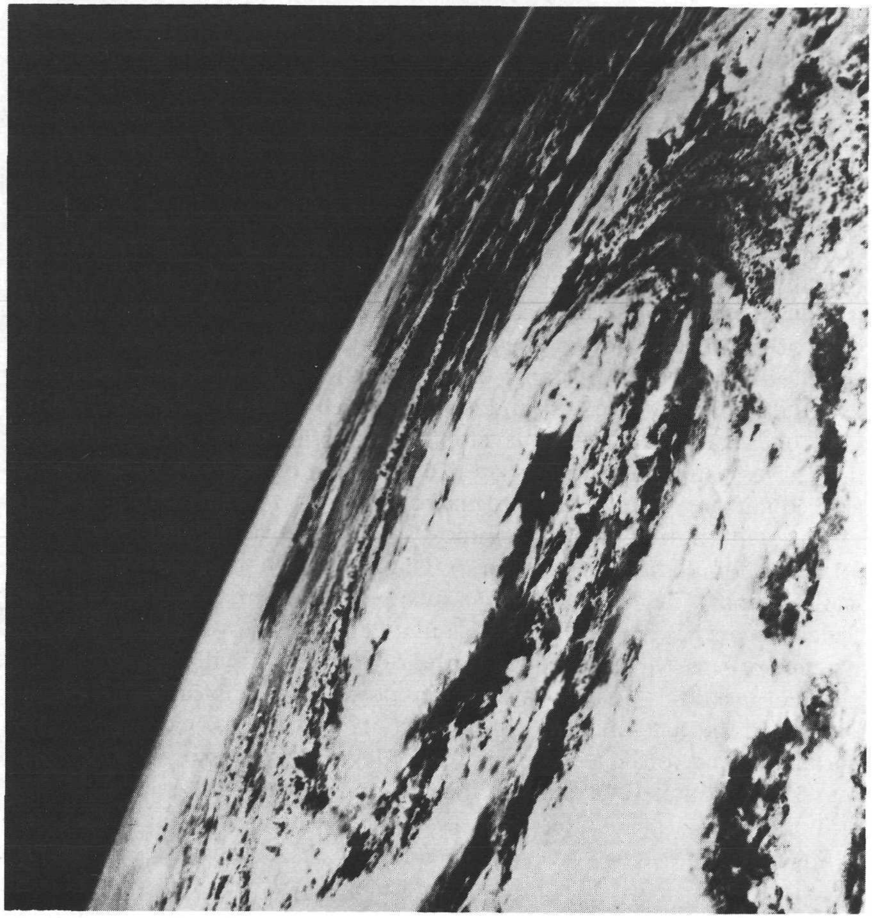


FIGURE IX.8. A hurricane viewed from space. (Photograph courtesy of NASA.)

and growth of hurricanes by modifying the clouds near the eye. Whether the state of the art is sufficient to make such tampering with dangerous systems productive is debatable; it is clear, however, that progress toward theoretical understanding is unacceptably slow and deserves a greater input of scientific and technological competence than the subject has hitherto received.

Another major scientific goal for meteorology is the control or modification of the weather. The most promising path lies through the modification of clouds and precipitation, because meteorologists understand the fundamental physics of these phenomena. However, the logistical problem is severe, and, so far, efforts have been limited to the use of silver iodide as a freezing nucleus for supercooled water clouds.

In recent years, it has been established that orographic clouds, layer clouds, and cumulus clouds can be modified by silver iodide seeding. However, the results are not always as expected—the existence of downwind rain "shadows," for example, suggests that knowledge of the interaction of physics and dynamics in clouds is far from complete.

The problem of climatic change is one of the most stimulating and difficult in atmospheric studies, and it has received added impetus in recent years as the question of inadvertent changes through the emission of industrial carbon dioxide and aerosol was raised. Until the present time, work in this field has been no more than speculation. There is no understanding of why the ice ages and lesser climatic changes occur.

Advances in atmospheric dynamics have little direct application to the problem of climatic change: Clearly, it is senseless as well as impracticable to integrate dynamic equations over centuries. Methods of computing average states of rotating, unstable systems are needed, and in the last two years the first suggestions of appropriate techniques have been made. At present, these suggestions do not include interactions with the oceans, and realistic solutions must include a model of the oceans and of air–sea interaction. The first convincing solution of a climatological steady state will be a landmark in the history of meteorology.

The interaction of the atmosphere, the biosphere, and the liquid and solid surfaces involves many different major problems, little understood but of great significance. Concentrations of minor constituents such as nitrous oxide and carbon monoxide are two problems that relate to questions of environmental quality. Of fundamental interest is the origin of the atmosphere and the chemical cycles that maintain it. This problem has received impetus from studies of the atmospheres of other planets. Those of Mars and Venus are almost pure CO_2, and yet their origin probably did not differ greatly from that of the earth. The feasibility of a runaway greenhouse effect has been raised as one important difference between Venus and earth. The probable absence of life on Mars and Venus is another.

Comparative studies of planetary atmospheres affect progress in many areas. The circulation of the Venus atmosphere will be investigated further because the atmosphere differs from earth's in its density, its chemical composition, the small Coriolis forces, and the uniform cloud cover. Mars differs from earth in other major respects. The investigation of these differences depends on the continuance of NASA's program of planetary exploration; if this program continues, these studies will remain among the most stimulating in the atmospheric sciences.

The fundamental questions of physical oceanography concern the conserved quantities and their processes of conservation. Scientists inquire into the sources, sinks, transport, and transfer of mass, momentum, vorticity, energy, and various chemical substances in order to rationalize and explain the water motions, stirring and mixing processes, and the mutual exchange of momentum, water, gases, and material substances between the ocean and the atmosphere.

New instruments are needed to provide a three-dimensional picture of the fields of velocity, temperature, and salinity. These data will allow a clearer view of the physical properties of the ocean interior. In addition, through the determination of the three-dimensional distribution of radioisotopes and trace elements, the interaction between physical and biochemical processes can be made clear.

Adequate theories of ocean processes are few. For example, a general theory of turbulence has so far been precluded by the complexity of the field equations, which admit solution only in trivial or highly idealized cases. Although advances have been made in the statistical theory of homogeneous isotropic turbulence, the present evidence is that the ocean is only *intermittently turbulent*, in a range between laminar flow and fully developed turbulence. Since nonlinearity and time dependence are of major importance, finite amplitude instability calculations are appropriate in this range of parameters; yet few relevant calculations have been made on the myriad of possible destabilizing modes inherent in the ocean. Moreover, the large variety of scales and energy sources present in the ocean allow a similar variety of balances in the conservation equations.

Paradoxically enough, scientists have a better knowledge of the larger scales than they do of the smaller scales. Although they have developed a broad average picture of the large-scale velocity, temperature, and salinity fields at the ocean surface, the vertical distribution of properties has been inadequately sampled. A sufficiently high-resolution picture of the vertical has only recently emerged, showing a rich microstructure, often remarkably regular. The existence of sharp gradients has important implications for the vertical mixing of water properties (see, for example, Figure IX.9).

Underlying the explanation of the microstructure is the physics of dissipative processes, which plays an important role not only in the small-scale flows but even in the largest-scale motions, tides, and currents. Turbulence measurements in the ocean, which might shed some light on these matters, are practically nonexistent. Other small-scale features, similar to atmospheric fronts, have been observed in the ocean but remain unexplained.

Although many of the broad features of the large-scale ocean circulation have been qualitatively explained by theory, troublesome observations and theoretical difficulties cast doubt on present understanding. Observations to test the theory have revealed small-scale eddies whose role in the vortic-

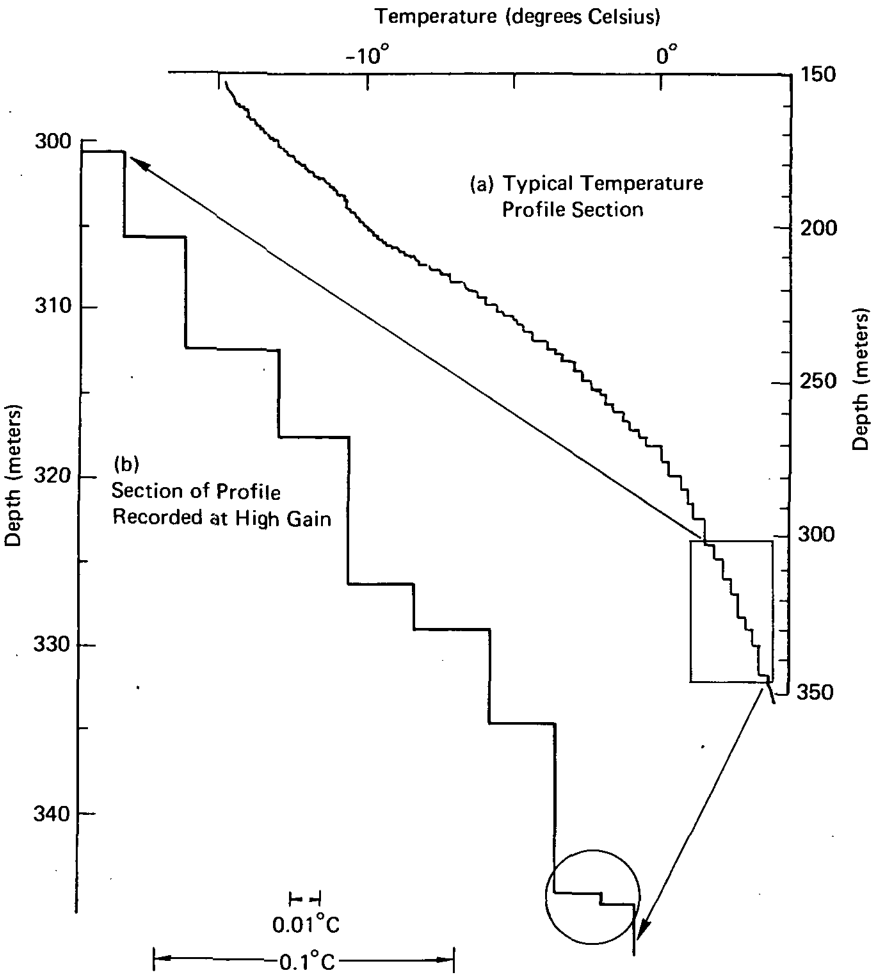


FIGURE IX.9. Thermal stratification observed in the Arctic Ocean.

ity balance of the ocean is unknown. Whether they dominate or are only passive is the central concern of a series of field observations and theoretical calculations presently in progress. The energy source of these eddies is unknown; it may lie in the eddies shed from major ocean currents, the baroclinic instability of the open ocean, transient surface forcing, nonlinear transport of energy from higher-frequency inertial motions, or internal waves and tides.

Although significant advances have been made in the subject of surface waves in recent years, field and laboratory experiments have shown that the accepted theories are not adequate to explain the wave-generation process. The subject of internal waves has received less attention. Theories applied to observation here have remained peripheral and speculative because of the lack of definitive field observations and data ambiguities. The microstructure mentioned above contaminates the high-frequency end of the spectrum; directional information requires arrays of instruments. Few observations have given any direct evidence that bears on the physics of internal wave generation, especially in the deep water. The recent development of deep-sea pressure gauges is one step toward the measurement of the statistical energy fluxes required for an explanation of the energy source.

Long-term measurements with recording current meters have revealed the ubiquitous existence of time-dependent motion at the inertial frequency, but the bandwidth, spatial coherence, intermittency, and source of these oscillatory currents is still poorly understood. Although in the immediate surface layers of the ocean, the observed inertial oscillations are probably locally wind-driven, the source of energy of the oscillations occurring at greater depths remains a mystery. Modes with periods longer than the inertial period have remained comparatively inaccessible to observation, but new middepth floating and bottom-mounted instruments are being developed in order to shed some light here. The nonlinear interaction of these longer period modes may be of fundamental importance to the physics of the large-scale general circulation.

Work on air–sea interaction has expanded in recent years with the use of large-scale arrays and aircraft measurements. Together with direct measurements of the covariances between the vertical velocity fluctuations and the property being exchanged, the large-scale measurements allow a more satisfactory understanding of the exchange processes and the development of parameters of the exchanges in terms of variables that are synoptically available. There is evidence that departures of the sea surface temperature from seasonal norms can affect the entire circulation pattern of the atmosphere, and it appears that the atmosphere and ocean system can be in a number of quasistable climatic states that shift from one to another when a particularly large perturbation occurs. How these anomalies arise and maintain themselves is not yet clear.

Scientific questions related to the planets and to interplenetary space have been mentioned a number of times in this chapter. We have not attempted to distinguish them from terrestrial research because there is no longer any important distinction. Earth will remain the only planet observed in detail, but the goals of enquiry are common to all accessible objects in the solar system. The only essential differences in planetary research are the high cost and the nature of the research tools. However, there is a difference of emphasis. Two of the most important and stimulating questions facing the physical and biological sciences are the formation of the solar system and the origin of life; planetary investigations are likely to yield their most spectacular results in these fields.

3 Applications to Man's Needs

Although the earth and space sciences have an important element of aesthetic and intellectual appeal as an area of human endeavor, the level of activity and support is largely a consequence of their direct applicability to man's needs. The aesthetic and practical aspects are not mutually exclusive; indeed, it is often difficult to distinguish between them. Problems of a fundamental character frequently emerge from empirical investigations directed toward short-term practical results. Conversely, almost any increase in present understanding of fundamentals in the oceans, atmosphere, and solid earth can be, and usually is, rapidly exploited. Natural phenomena must be understood before they can be exploited, forecast, modified, or controlled. For example, in the realm of solid-earth geophysics, the concepts of sea-floor spreading and global tectonics currently are being used as a guide in the location of previously overlooked petroleum and mineral deposits. The discovery of the importance of fluid injection is being used in preliminary attempts to control earthquakes. Routine compilation of earthquake locations for fundamental problems in earth physics is critical in siting nuclear reactors.

In practice, it is not difficult to distinguish between investigations suitable for universities and academically oriented research institutions and those best undertaken by federal agencies, industry, or some large private research organizations. This division of labor usually relates to the size and administrative complexity of the project, whether results are required on a relatively short time schedule (anything less than five years is obviously difficult for a university) and whether long-term continuation can be anticipated so that senior research staff can be employed under satisfactory conditions.

Probably, the earth and space sciences should be regarded as a continuum embracing the entire range from pure to applied physics research. The object of this chapter is to point to some of the final products of this process.

3.1 OCEANIC PHYSICS

The food supply man derives from the sea is a major stimulus to the study of physical oceanography. The relationships between marine biology and physical oceanography are much closer than the corresponding relationships between life sciences and atmospheric sciences. The density of water is so nearly that of organisms and the interaction between the physical processes in the sea and marine life is so close that an intimate relationship exists even at a very basic level.

Maritime commerce depends heavily on oceanographic knowledge. For example, some shipping routes are being programmed to minimize travel time and fuel consumption by allowing for waves and currents. The U.S. Navy also uses oceanographic information in this way and is engaged in many other aspects of oceanography. Problems of propagation and scattering of underwater sound, which depend on the distribution of temperature and salinity in the sea as well as on bottom properties and scattering organisms, are of major interest.

The production of oil on continental shelves and the mining of the sea floor present oceanographic problems. Offshore oil production, now a burgeoning industry, requires many kinds of oceanographic data, particularly wave and storm predictions.

Coastal engineering, including the construction of harbors and recreation facilities, also offers many challenges to the oceanographer. In general, the failure rate of projects has been unacceptably high; breakwaters are destroyed and jetties lead to either deposition or erosion of adjoining beaches. A better understanding of the interaction of oceanographic processes with nearshore sediments is essential to the development of coastal engineering.

Weather and climate prediction require oceanographic input. The melting of northern hemisphere glaciers before 1940 was associated with a rise in sea level of the order of 1 ft per century. Such changes could have far-reaching long-term effects. Problems of marine and coastal safety from tsunamis and tidal waves as well as from sea and swell are other critical areas of concern that demand both meteorological and oceanographic knowledge.

Two examples of the application of meteorology and oceanography to important environmental problems have recently appeared in the literature. The first concerns the problem of storm surges at Venice. In the Adriatic Sea, storm surges that endanger the foundations and treasures of this old city occur frequently. The damaging effects of the surges increase every year as the city slowly sinks. The solution seems to be to close the lagoon's three entrances against high water during a surge but not to interfere at other times with the free movement of ships and of the tides that flush the lagoon clear of sewage twice daily. The main theoretical problem is to devise an accurate scheme to predict the surges at least 6 to 8 hours in advance from information on atmospheric pressure, winds, sea level variation, and local and bottom topography. Theoretical calculations show that wind stress over the northern part of the Adriatic is the major contributor to the forcing of the surges that appear at Venice and that the 22-hour seiche can remain in the Sea several days after its generation by a storm. When new storms are in phase with the remaining seiche and the astronomical tide, the surges are greatest. Rotation, friction, and nonlinearities play only a small role in the development of the seiches. A theoretical model of the surges has been devised that allows consistently good predictions of the sea level at Venice 6 to 8 hours in advance.

A second example is the prediction of the position of warm currents, such as the Kuroshio Current, a major feature of the general circulation of the Pacific Ocean. The position of such currents affects such industries as fishing and agriculture. Recently, there has been progress toward the understanding and eventual prediction of the meander phenomenon.

3.2 SOLID-EARTH GEOPHYSICS

Acquiring a full understanding of earthquakes and their effects is one of the high-priority tasks for geophysical research in the next decade. Major objectives are to detect and monitor the stress buildup preceding a major earthquake and to anticipate its effects on structures and public works. Much work concerned with the detailed study of local geology and geophysics has made possible the prediction of the response of the ground to an earthquake disturbance. A current issue in public policy is the inclusion of such considerations in zoning and regional planning. The possibility of artificial modification of stress buildup in the crust is less well understood and is receiving intensive study. A full understanding of the role of underground water in stressed regions should make the engineering of gradual stress release feasible.

The understanding, and possible forecasting, of earthquake-generated tsunamis, or seismic sea waves, involves seismology, oceanography, and meteorology. For example, only certain types of earthquakes and only earthquakes in certain places can generate tsunamis. This information can be supplied quickly by a continuously manned seismic observatory. Earthquakes that will generate tsunamis also generate acoustic-gravity waves in the atmosphere that travel faster than the tsunami itself. Thus, an early warning system is possible, with a low rate of false alarms.

In petroleum exploration, advanced geophysical methods combined with miniaturized instrumentation and computers have greatly increased the resolution of mapping the vertical dimension. With the discovery of oil under the ocean floor, marine engineering has become an important new area of application, as have the location and extraction of minerals using electromagnetic methods of exploration (resistivity, induction, aeromagnetics, and the like).

The storage and retrieval of groundwater is a critical problem. Studies of the physics of flow in porous media and field investigations to determine relevant parameters for real systems are leading to the use of underground storage as a conserved resource. The transport of pollutants by groundwater is an equally important facet of this problem. The theory of groundwater flow has implications for the most economic means of extracting oil.

Geophysical methods (resistivity and seismic techniques), in addition to more conventional drilling, are used to locate groundwater. The effect of near-surface and deeper groundwaters on near-surface rocks is to modify their mechanical properties; the application to dam safety and slope stability is obvious.

Geothermal power is an economic possibility in many areas. The large-scale exploitation of this resource depends on a full program of research in drilling and site engineering problems and the detection of geothermal fields.

Society depends heavily on numerous rare elements from the crust—helium, silver, copper, tungsten, and the like. Man is now forced to plan carefully to avoid depletion of these resources. Possible techniques include recycling, substitution, and lower-cost extraction. The location and assaying of the world resources of these rare elements are integral parts of this effort and primary tasks of the geological sciences. Geophysical insight, as well as geophysical techniques, is important in identifying metallogenic provinces. However, commercial mineral prospecting is too short-range in its objectives to define the extent of the earth's mineral reserves adequately. What is needed is a much greater investment in research that is intermediate between commercial prospecting and basic research in geology and geophysics.

3.3 LOWER-ATMOSPHERE PHYSICS

The social and economic implications of lower-atmosphere studies are the most familiar examples of the usefulness of physical studies of the environment. Obviously, the lower atmosphere is directly relevant to man's well-being since it provides the air he breathes and the storms he endures. It is the most valuable natural resource—the one most susceptible to destruction by man's activities and the one most amenable to change for the better by his intervention. It is, at the same time, both ally and enemy. Short-term forecasting and long-term planning can mitigate its dangers—severe storms on the one hand and pollution on the other. The popularity of television weather programs is sufficient testimony to the importance of weather forecasting for recreational purposes and other aspects of daily life. In agriculture, good forecasts can help to conserve labor and materials and improve harvests. Air transportation could not operate effectively without meteorological services. Ships can avoid hazards when provided with accurate predictions. Even road traffic needs warning of severe conditions. The construction industry requires forecasts for scheduling of work forces, materials, and equipment. Effective water management and conservation practices depend on information on floods and droughts. Public utilities operate more efficiently with adequate warning of adverse weather conditions. The overall value of extended and reliable weather forecasts has not been accurately assessed, but it is generally believed that possible savings greatly exceed expenditures on research and operations.

Although weather prediction has long been recognized as an especially complex application of the principles of fluid dynamics and thermodynamics, it has, until recently, been practiced largely as an art. However, an essentially complete mathematical prediction model was formulated by Richardson[3] some 50 years ago. Observational and computations problems encountered at that time have now been overcome, and numerical predictions of the distributions of weather variables are made routinely by the National Meteorological Center of the National Oceanic and Atmospheric Administration (NOAA) for periods of up to three days.

The outstanding limitations on current numerical prediction models are due to (a) gross approximations made in representing the transfer of energy, momentum, and water vapor within the boundary layer; (b) difficulties of handling change of phase in clouds; (c) the inherent inability of the models to represent explicitly atmospheric features smaller than the grid scale; and (d) the lack of global data. The World Weather Watch (WWW) and Global Atmospheric Research Program (GARP) have been launched under international auspices to extend the range of useful prediction as nearly as possible to the "theoretical limit" for deterministic general circulation models (about two weeks). This limit is based on the results of numerical experiments that show that in about two weeks of forecast time, small initial errors in subgrid scale features may grow until their effects are as large as the difference between randomly selected natural states of the system.[4]

The attempt to extend the range of prediction is an experiment on a grand scale. It is being conducted in several stages that will extend through the 1970's, but significant, measurable improvements already have occurred in surface and upper-atmosphere forecasts.[5] The 36-hour "skill score," which is a rough measure of the root-mean-square vector error of the wind forecast, has decreased as each major model improvement has been introduced. The model in use since 1966 represents, at least crudely, boundary layer and cloud processes; with it, the routine period of prediction has been extended to 72 hours. New developments of storms or of high-pressure regions can now be predicted with considerably greater confidence than was possible a few years ago. Perhaps the best evidence of the improved quality of the numerical predictions is the fact that experienced forecasters, who are characteristically proud of their individual skill and judgment, now rely on the numerical prediction to provide the basis for their forecasts.

Much of the variability of the weather (roughly half, on the average) is contained in subgrid-scale features; these features are predicted by subjective methods based largely on accumulated experience and individual judgment. This small-scale variability is superimposed on the variability of the numerical prediction and therefore must degrade the model forecasts for specific small areas and specific times. The limited comparisons that are available indicate a gradual improvement extending over several decades in forecast accuracy for local conditions.[6]

Less familiar but of great potential importance is the possibility of weather and climate control. A report of the National Academy of Sciences[7] outlines the many possibilities for small- and large-scale modifications and gives evidence of some apparent successes. It also finds that any real advance depends on a fundamental understanding of a variety of weather phenomena. Although this understanding does not yet exist, it is the target of many research programs and, for good or ill, man may have substantial control over some aspects of the weather in this decade.

Climate control can be inadvertent as well as intentional. Such issues relate essentially to quality of the environment and are discussed in a subsequent section.

3.4 UPPER-ATMOSPHERE AND INTERPLANETARY PLASMA PHYSICS

The more remote the events from the earth's surface, the less direct is their impact on man's material needs. Clearly, upper-atmosphere and plasma studies do not have the same degree of immediacy as those pertaining to the solid earth, the oceans, and the lower atmosphere. Nevertheless, applications are a major stimulus to these studies.

Studies of the upper atmosphere grew largely from a desire to understand radio propagation and reflection of radio signals from the ionosphere.

Developments in monitoring and prediction of ionospheric processes will continue to be important in the coming decades. Understanding of the interplanetary plasma is an essential feature of ionospheric prediction and has great potential significance for the viability of man in the space environment.

The recent congressional debate on funds for the supersonic transport (SST) showed the importance of preserving major competence in aeronomy. The possibility of changes in atmospheric ozone as a consequence of catalytic reactions involving aircraft exhaust gases became a major issue. It was suggested that operation of an economically viable SST fleet would lead to a modest decrease (about 1 percent) in atmospheric ozone, which would result in an increased surface irradiation below 3100 Å. Estimates suggested that such a change in ozone could result in as many as 10,000 new cases of skin cancer per year in the United States alone.

The general features of the oxygen–ozone chemical chain were first described by Chapman some 40 years ago. The chemistry is complicated somewhat by the addition of H_2O; these complications were first discussed by Bates and Nicolet some 20 years ago. As a result of reactions with OH, H, and HO_2, derived from water from jet engines, there will be some reduction in the O_3 level. Unfortunately, a precise estimate is not possible at this time since the crucial chemical reactions have not been studied quantitatively and their rate coefficients are essentially unknown.

The effects of other pollutants have been even less carefully examined. For example, there are potentially serious consequences of addition of NO_x. The natural NO_x level in the stratosphere is unknown, although mixing ratios are available at higher (80 km) and lower levels. If one assumes that the natural stratospheric level can be estimated by linear interpolation between these mixing ratios, one would conclude that the SST will represent only a minor perturbation of the natural system. However, this reasoning, which was widely accepted in the recent debate, is prone to error. The atmosphere acquires NO_x from two principal sources, one the result of microbial activity at the surface, the other a consequence of ionospheric chemistry. The SST source is larger than the natural ionospheric source by approximately a factor of 10. Thus, in principle at least, the SST could markedly alter the natural NO_x concentration, with possible effects not only on the ionosphere but also on ozone.

Another aspect of aeronomy and interplanetary studies should not be minimized. A strong feeling exists in the atmospheric science community that, in the long run, the atmosphere must be considered as a single system. Even those most devoted to lower-atmosphere studies do not find it meaningful to draw a line in the atmosphere below which it is considered useful and above which it is of academic importance. The changing structure of many universities reflects this outlook; they are attempting to integrate aeronomy and lower-atmosphere studies, to the benefit of both.

3.5 LUNAR AND PLANETARY PHYSICS

Lunar and planetary research once appeared remote from man's needs, but, as the space program has developed, this view has changed. Meteorologists, aeronomers, and solid-earth geophysicists are increasingly interested in the moon and planets, since new knowledge of these objects has stimulated new thinking and added significantly to the knowledge of the earth. This new thinking relates more to the scientific goals of Chapter 3, but a brief discussion will help to emphasize the close relationship between pure and applied research in the earth sciences.

The Space Science Board of the National Academy of Sciences–National Research Council[8] gave as one of the three major goals of the planetary program "to provide for progress in our understanding of the dynamic processes that shape man's terrestrial environment." The reason for this emphasis is that problems of the solid earth and lower atmosphere are characterized by the extreme complication of their many interacting processes and the paucity of essential data. To make progress it is usually necessary to accept an observational definition of much of the system under study and to attempt an explanation of other features from these given conditions. Such procedures are essential, although they often obscure the fundamental physical processes. What needs to be explained and understood is too often accepted as a given feature of the physical world. New data on the planets, referring to systems similar in major respects but widely different in details, have led to new ideas and a re-examination of fundamentals that have been extremely stimulating to terrestrial research.

It is instructive to compare the atmospheres of Mars, Venus, and the earth. On Mars, the low atmospheric density, low water-vapor concentrations, and paucity of clouds indicate that the atmosphere is close to radiative equilibrium and that dynamical effects on the temperature distribution are small. On Venus, the very large atmospheric densities and dense cloud cover act to minimize the effects of radiation so the temperature field is expected to be strongly controlled by the wind systems. On earth, radiative and dynamical influences are more nearly of equal importance. Clearly, fundamental gains should result from studies of all the atmospheric systems provided by nature. Such studies present meteorological theory with different sets of circumstances from those afforded by the earth. All must be explicable by a correct and general theory; therefore, comprehension of earth's atmosphere can be improved in the process.

The slow rotation of Venus means that Coriolis forces will be small, a situation found on earth in the tropics, where the significant part—the vertical component of the rotation vector—is also small. Thus the circulation of Venus may be fundamentally related to that in the tropical region of earth, an area less well understood than the middle and high latitudes. Another potential relationship to terrestrial motions is to the part of oceanic circulation that is thermally driven. The oceans absorb solar radiation close to the top boundary. Much of the deposition of solar energy on Venus could be in the upper part of the clouds, and a parallel to terrestrial oceans exists at least in some theoretical models.

Hypotheses regarding the origin and evolution of the earth, including dynamic aspects such as tectonism, sea-floor spreading, and maintenance of the magnetic field, probably will be dramatically affected as the study of other planets continues. Surface processes erased most evidence of the early history of the earth. The study of the surfaces of the moon and, possibly, Mars will give information that extends back to the origin of the solar system. Studies of the chemistry of atmosphere of Mars and Venus, perhaps unmodified by biological activity, may aid in understanding the early evolution of the earth's interior and its atmosphere. The study of the interiors of other planets, including tectonic activity, will advance understanding of the evolution of the earth's core, mantle, and crust and perhaps provide clues to the tectonic driving mechanism. Ideas about the terrestrial dynamo-magnetic field already have been affected by the discovery that Mars and Venus do not have such a field.

3.6 ENVIRONMENTAL QUALITY

The current worldwide concern for the quality of the environment is not new to the meteorologist or the oceanographer. Knowledge of small-scale diffusion and transport by planetary-scale motions in the atmosphere and the oceans has advanced steadily since World War I, and each advance has found applications: to the diffusion of effluent from smokestacks; to the dispersal of radioactivity in the stratosphere; to the trapping of airborne pollutants in urban areas; and the like. The problems involved are not unique to studies of environmental quality but are relevant to such central themes as the diffusion and chemical reactions of minor constituents; the nature of planetary boundary layers; and local and global transport of water vapor, heat, and momentum.

Because of these factors, the earth sciences have been able to make a relatively rapid response to the popular call for increased environmental research. In a recent study by the National Research Council's Committee on Atmosphere Sciences,[9] air quality was identified as one of the three major areas of meteorology in which rapid advances could and should be made in the next decade. The President's Commission on Marine Sciences, Engineering and Resources[10] could point to specific increases in scientific and engineering activity at the rate of $65 million per year to meet the needs of coastal zone management. In the recent debate on the SST, aeronomers were able to make immediate (if somewhat belated) contributions on key issues.

The position of a conventional physicist with respect to environmental problems differs greatly from that of an atmospheric or marine scientist. He is sometimes told that his discoveries, together with derived technology, are partly responsible for the current threat to the environment and the quality of life. This view has been one cause of a spreading apathy toward the physical sciences, particularly among young people. The earth sciences, with their relationship to both physics and environmental quality concerns, probably provide the best bridge between the two areas.

An increasing number of young physicists are expressing a desire to work in environmental research. Even with adequate motivation, however, they encounter difficulties in the direct use of a conventional education in physics. Physicists who have made the transition to environmental areas frequently express the opinion that the content and attitudes of frontier physics are deficient in regard to applications to human needs.

Areas of environmental research exist in which a problem can be isolated in terms familiar to the physicist. Many of these areas involve new techniques of measurement and new instrumentation. The development of new spectral imaging devices for aircraft and satellites, particularly in the infrared, and the interpretation of the results in terms of the physical, chemical, and biochemical properties of the earth's surface currently claim a great deal of attention. Laboratory measurements and theoretical calculations of reaction rates and absorption cross sections are continually in demand. Laser sounding of the atmosphere will probably be an important diagnostic tool in future decades. Radioactive and other trace elements can be used to study atmospheric diffusion and circulation. Extremely sensitive strain gauges will be needed for earthquake monitoring and prediction. Much progress in instrumentation will be required before completely satisfactory measurements of almost any physical parameters at great depths in the oceans will be possible.

To these and many other problems the physicist (particularly the experimentalist) can make a direct contribution. However, his role is not one of leadership, nor does it involve the formulation of new research directions. Consequently, it does not exploit to the fullest the powerful methods of the physical sciences. The potential value of a physicist outside the conventional core subjects of physics goes beyond the development of techniques. A physicist need never write down Maxwell's equations nor operate a particle counter, but he can still contribute a style of research, an open-mindedness, and a confidence in man's ability to solve problems. He can offer a confidence that difficulties can be overcome; scant respect for conventional wisdom; a willingness to discard shopworn concepts; the ability to quantify, first with order-of-magnitude estimates, then, if necessary, with numerical or sophisticated analytical techniques; and the ability to understand the relative importance of different data and different ideas.

A strong motivation to participate in environmental research, even when coupled to the above advantages, may not suffice to produce an effective contributor. Some knowledge of facts in the earth sciences is necessary to form a useful judgment in this complicated area. Observational methods must be understood. The ability to work with other disciplines toward a common objective must be learned. A combination of ability and naïveté can sometimes produce spectacular results, but this is not the way for most research workers; if physicists are to be effective in environmental research, deliberate modifications in physics educational programs will be required.

The earth sciences represent the easiest and most familiar route for a conventionally educated physicist to enter the environmental area; therefore, the educational problems discussed in Chapter 4 are relevant in the wider context of broadening and utilizing a traditional physics education so that physicists can be more effective in the environmental sciences.

4 Education

"It is often implied that contemporary graduate and postdoctoral training is becoming so narrow that students have lost the traditional breadth of outlook and flexibility expected of a physicist." This statement in the charge to the Physics Survey Panels describes one of the main problems in current physics education. Particularly at the graduate level, physics education during the past 25 years has largely become training for research in the core subjects of physics rather than a general education to tackle the world's novel problems. During the same period, education in engineering and earth sciences has been moving with considerable success away from a narrow professional approach toward one increasingly involved with all the natural and social sciences. Throughout this interval, however, physics departments have continued to attract some of the most talented students. A pressing need exists to guide some of the excellent students available to the physics departments into other activities in physical science.

The traditional pre-eminence of physics and its curriculum among the sciences is the result of an education that gives the physicist the ability to tackle difficult and novel problems. For the past 30 years, the class of problems that brought the physics establishment to the status of a major national resource has been related largely to missions of various agencies of the federal government. It is increasingly evident that the future will require a major effort in many aspects of the large-scale physics of the environment that previously were regarded as applied physics or engineering. The national investment in these problems requires a remodeling of science education and offers a major challenge to physics departments. The values of physics must be broadened to attract some of the best students into these new activities of wider scope, while at the same time continuing the time-honored quest for fundamental knowledge at the largest and smallest scales.

4.1 DEVELOPMENT OF EARTH AND PLANETARY PHYSICS

Earth and planetary physics evolved gradually over a long period of time through the development of a number of disciplines that initially were regarded as virtually independent. These include meteorology, aeronomy, geology, geodesy, oceanography, hydrology, lunar astronomy, planetary astronomy, and solar astronomy.

During the past 25 years, interest in the physics of atmospheres developed rapidly and has exceeded all expectations. The advances closely followed several decades of less dramatic but substantial growth in atmospheric science. At the beginning of the century, the earth's atmosphere implied mainly the region below the tropopause; it was vaguely presumed that above this region diffusive equilibrium existed at constant temperature. Geomagnetic variations had given a slight indication of the possible existence of an ionosphere, and Marconi's trans-Atlantic radio experiments had not yet been performed. Almost nothing was known about the complex weather phenomena of the lower atmosphere; even a comprehensive description of these events was impossible with the limited communications available. In the ensuing half century much of the present knowledge of the earth's atmosphere was accumulated. Today, the solar atmosphere, the interplanetary medium, and the planetary atmosphere are considered parts of a single vast fluid system with electrodynamical and magnetodynamical properties that can be explained, in principle, in terms of the fundamental laws of physics.

What is now recognized as a single field, atmospheric physics, developed piecemeal through the efforts of many scientists in separate university departments, attending different scientific meetings, and publishing in different journals. Study of the earth's lower atmosphere took place largely in meteorology departments, occasionally within engineering schools. In the United States, the study of the ionized upper atmosphere of the earth was conducted for the most part in electrical engineering departments using radio techniques. Astronomy departments, using optical techniques, contributed to the development of the solar atmosphere. Recently, however, these departments have adopted and applied the methods of radio astronomy, which were introduced in electrical engineering departments. Until the last decade, planetary atmospheres other than the earth's can scarcely be said to have been systematically studied on a nationwide basis; instead, such studies were the pastime of a few lonely individuals, usually affiliated with astronomy. The study of cosmic rays, and more recently of Van Allen particles, occurred principally in physics departments.

In much the same way, solid-earth geophysics grew from the concern of a few seismology observatories into a broad-scale effort that requires people educated in a wide range of physics subfields. Current research into the state and dynamics of the earth involves disciplines such as theoretical and experimental solid-state physics, classical mechanics, wave propagation, solid mechanics, and ultrastable field instrumentation. Leadership in providing broad education in these fields for application to the solid earth has come from a few institutions that established autonomous programs within geology or planetary science departments. Physics departments in this country generally have not taken an active part in these programs. In contrast, in British, Canadian, and Australian universities, a major fraction of the significant work in solid-earth geophysics has taken place in physics departments, and the much smaller total program in these countries has a disproportionately large impact.

Some of the major work in oceanography also is conducted in physics departments, for example, at Cambridge University. However, the dependence on extensive seagoing facilities led in most countries to the establishment of somewhat remote institutes in which the teaching benefited little from developments in related subfields of physics. This situation changed with the development of the study of underwater sound propagation, and the relation between physics and physical oceanography has grown closer in recent years. The need for a closer link with physics and other basic sciences was a major factor in the formation of the University of California, San Diego, contiguous to the Scripps Institution of Oceanography. Some other institutes moved their major activities from the field station to the home campus (for example, Johns Hopkins University). The Woods Hole Oceanographic Institution made arrangements with the

Massachusetts Institute of Technology and Harvard for a closely coordinated teaching program.

Universities now recognize that activities that developed piecemeal in physics, geophysics, astronomy, geology, meteorology, and electrical engineering departments are merging and evolving into integrated broad fields of endeavor. This trend makes obsolete what was once a reasonable academic structure. University efforts to cope with new fields emerging through the fusion of established disciplines and subfields are particularly hampered by the absence of a coherent, university-wide approach to all the physical sciences.

4.2 THE CONCEPT OF EARTH AND PLANETARY PHYSICS

The circulation of the atmosphere and transports in it cannot be separated entirely. Both atmospheres and oceans were formed as the result of geological processes and now constitute a major modifying factor for geomorphological processes. This commonality of subject matter and a similarity of research methods suggest that the study of planets should encompass the solid, liquid, gaseous, and plasma states, viewed as parts of a single system. Table IX.1 depicts the interrelations among the various components of what could be described as earth and planetary physics. Studies are arranged in horizontal rows, depending on the basis of their association with (a) atmospheric phenomena, (b) surface phenomena, or (c) interior phenomena. Arrangement in the vertical columns depends on the locus of the phenomena with which a study deals, that is, whether they pertain to the earth in particular, the planets in general, or the sun.

A desirable university organization would be one that gives appropriate recognition to earth and planetary physics as an integrated discipline and

TABLE IX.1 Earth and Planetary Physics

Types of Phenomena	Locus of Phenomena		
	Terrestrial	Planetary	Solar
Atmospheric physics	Meteorology and aeronomy	Planetary atmospheres	Solar wind, solar atmosphere, and photosphere
Surface physics	Geology, geodesy, oceanography, and hydrology	Planetary and lunar surfaces	–
Interior physics	Earth's interior	Planetary and lunar interiors	Solar energy and evolution of the solar system

provides for study and research based on both the horizontal and vertical cross sections depicted in Table IX.1. Such a department would present earth and planetary physics as a broadly conceived discipline capable of commanding the attention of both faculty and students of high intellectual ability.

4.3 THE SUPPLY OF GRADUATE STUDENTS

The objective of graduate education is to bring a student from the point at which he feels reasonable confidence in his ability to understand what other people have understood to the point at which he has the confidence to face novel situations. The principal way of achieving this educational objective on a university campus is research.

The need for an adequate supply of graduate students with an appropriate background for study and research in earth and planetary physics is great. To meet this need requires a program that will attract students with sound undergraduate training in mathematics, physics, and chemistry and stimulate them to study and do research in earth and planetary physics. If students were brought into contact with research in the earth and space sciences in the same way that they now encounter the research frontiers in the core subjects of physics, this goal might be achieved.

The need for graduate students with a good physics background is not restricted to earth and planetary physics. Much of the graduate work in materials science and quantum electronics, and much of that associated with nuclear reactors, is now conducted in departments other than physics, usually in engineering schools. Difficulty in obtaining students with adequate undergraduate preparation in physics is a common experience for many of these programs.

The earth sciences—geological, meteorological, and oceanographic—generally have not been a major choice of physics graduates. In 1970, only 3 percent of the physics graduates at the bachelor's or master's level went on to a PhD in the earth sciences. On the other hand, a significant fraction of earth science PhD's have been awarded to physics majors (Table IX.2).

The figures in Table IX.2 indicate an upward trend in the number of physics majors who conduct research in the earth sciences. This upward trend probably will continue, with an increase also in the total number of PhD's. The examples of engineering physics and biophysics provide comparisons with fields that traditionally involve a crossing of disciplinary lines. Engineering departments more than others have been able to expand basic physics training to keep pace with their own doctoral demands.

TABLE IX.2 Percentage of PhD Recipients Whose Undergraduate Major Was Physics

Discipline	PhD's per Year	Year 1962	Year 1969
Geophysics	56	29%	32%
Meteorology	50	20	31
Oceanography	53	8	15
Engineering physics	75	22	21
Biophysics	107	35	49

Experience indicates that physics majors who pursue graduate work in disciplines other than physics are even more likely to achieve the PhD than those who remain in physics. Between 16 percent and 19 percent of the physics baccalaureates typically plan graduate work outside physics, and the number is likely to be about 25 percent of the 1971 graduates. Of those who seek the PhD in disciplines other than physics, 85 percent are successful. Clearly, encouraging physics majors to pursue graduate work outside the conventional areas of physics is a highly responsible counseling procedure.

What are the universities doing to meet the demand outside the physics departments for students with good fundamental training in physics? Unfortunately, there is little to report. Physics graduate schools are interested primarily in research on those aspects of physical science in which the laws of physics are imperfectly understood, and the present arrangements for university undergraduate education in physics provide a more than sufficient supply of students for such research. However, these arrangements do not fulfill the requirement in many other branches of the physical sciences in a university for graduate students with good physics background. If undergraduate physics departments were to attempt to meet this broader requirement, the physics faculty would have to allocate far more time and attention to physics education than that which is necessary merely to ensure the success of their own graduate physics programs. Understandably, many are reluctant to make such a commitment.

The problem might be alleviated somewhat if professors of physics, particularly the younger ones, could be stimulated to spend a summer, a term, or a year devoted to the study of aspects of physics, particularly earth and planetary physics, not usually encountered in physics departments. An encouraging indication of the probable success of such a program is that earth and planetary physicists appear to be reasonably mobile in terms of research interest: Almost half of those in this subfield in 1970 worked in other physics subfields in 1968, as Table IX.3 shows.

TABLE IX.3 1970 Population of Earth and Planetary Physics in Terms of 1968 Subfield

1968 Subfield Source	1970 Earth and Planetary Physics (%)
Earth and planetary physics	55.6
Astronomy	9.5
Plasmas and fluids, atomic and molecular physics, optics	10.6
Elementary particles, nuclear physics, condensed matter	13.3
All other subfields	11.0

The intellectual mobility of the earth and planetary physics group is reflected in their range of secondary expertise. For example, when earth and planetary physicists were asked to list four specialties in which they were scientifically competent, about half of these responses were in other subfields of physics. In contrast, only 27 percent of the responses of solid-state physicists indicated expertise outside their subfield; 70 percent of physicists in biology specialties cited competence outside biology.

A program to provide opportunities for research and study in the earth sciences through a professional sojourn—a summer, a term, or a year—could increase the number of interested faculty in two ways: first, it could enhance further the already substantial migration from other subfields, and, second, it could sustain and deepen the interests of physicists in the earth sciences who might otherwise return to other areas of research.

4.4 UNIVERSITY ORGANIZATION

How should universities go about achieving the necessary emphasis on the teaching of undergraduate physics required to meet the needs of the entire physical science component of a university, and especially the needs of earth and planetary physics?

It would be possible to effect the necessary changes in the teaching of basic physical science by modifications in the programs of existing departments such as engineering physics, engineering science, meteorology, geology, astronomy, electrical engineering, nuclear engineering, aerospace engineering, engineering mechanics, and materials science. In view of the present academic structure of universities, this is the line of least resistance and to a limited extent has already been adopted. For example, undergraduate programs giving students a background in physics have been successfully operated in engineering physics departments in engineering schools at a number of universities. However, this development is not without its dangers. The universality and breadth of a physics education, its most important characteristics, could be sacrificed to the limited demands of the graduate school in which it is taught. Further, high school faculties might not be aware of and thus might not advise their better physics students to enter such programs.

An alternative and more satisfactory approach for the long term lies in an academic structure so arranged that all university faculty involved in fundamental research in the physical sciences play a full role in the teaching of basic physical science in some coordinated scheme in which the physics department would be the central, but not the sole, contributor. In this coordinated program, earth and planetary physicists should assume an active role, not only in the teaching of undergraduate physics courses but also in planning the content of curricula for undergraduate physics majors. These students tend to regard physics as the ensemble of interests of the faculty with whom they come into contact in the undergraduate physics major program. Through problems and examples in undergraduate physics courses, the graduate faculty in earth and planetary physics could foster awareness of the content and challenge of this subfield among undergraduate physical science students.

The Panel's study of manpower, reported in Chapter 8, suggests that the current university structure, which includes the university research institute, favors the research function of academic employment over the teaching function in earth and planetary physics. The distinction between university faculty and university scientist is real and has self-perpetuating elements. It is not easy for a scientist to undertake teaching after many years of concentrated effort in research, nor is the established faculty likely to welcome the researcher in preference to an experienced teacher. But, if the functional barrier between the research and teaching staff could be overcome, substantially increased participation of earth and planetary physicists in physics teaching could be achieved without increasing existing university resources.

Broadening the scope of physics is important for the future health of earth and planetary physics and other subfields and disciplines with similar problems. However, this Panel has not proposed specific steps to achieve this broadening, for many exist, and the choice will vary from institution to institution. Activities of the department of physics could be extended; a university-wide program for physics could be developed; or the many aspects of physics could be combined into a school of physics under a dean.

Whatever the solution, there are certain essential features: more students than are required for the core subjects of physics; basic courses taught by people from applied physics, geophysics, astrophysics, and the like—a deliberate selection representing the full scope of the subject; proximity of all physical research areas to the teaching facilities; and maximum fluidity in the early stages of graduate school so that students can seek thesis research topics far removed from their earlier interests.

We have concluded that the undergraduate educational problems of earth and planetary physics will not be solved satisfactorily by modifying the activities of departments of earth sciences or engineering science. The purpose of such departments is primarily to train professionals. Their accumulated data and records of phenomena are an essential resource for the earth and planetary physicist, but the departmental structure separated from physics is inappropriate for undergraduate education for research in earth and planetary physics.

If departments separated from physics are inappropriate at the undergraduate level, is there a place for graduate departments offering a PhD degree in earth and planetary physics? The answer depends on the nature of the departmental organization. If department implies a field, division, or set of regulations in a graduate school designed to facilitate a particular combination of study and research leading to a PhD with a thesis in earth and planetary physics, the answer is clearly yes. We further believe that such a course of study should emphasize the unity of earth and planetary physics, as depicted in Table IX.1, and should be strong in graduate courses in physics and mathematics.

If, however, a department is taken to mean an autonomous organization to which tenure faculty are appointed solely to direct research in geophysics, oceanography, meteorology, or space physics, we are less confident of its validity. Such a department will probably not have the visibility and glamour that would lead to an input of undergraduates of the caliber and education appropriate to the difficult research, and, ultimately, it would have to seek association with an undergraduate department or find itself largely divorced from the essential functions of a university. When it comes to the choice of the appropriate undergraduate department, the traditional undergraduate educational programs in a department of earth sciences are too weak in physics and mathematics to serve as adequate preparation for research in earth and planetary physics. Some participation in undergraduate teaching in physics is virtually a necessity.

The Panel's views on graduate teaching and research in earth and planetary physics can be briefly summarized as follows: A program of graduate teaching and research in this subfield must be based on a sound fundamental education in mathematics and physics. It will probably exist outside the traditional departmental structures. It may take the form of an interdepartmental program, or be associated with a research center or a research institute. Graduate teaching should stress the unity of studies of the entire solar system. The primary emphasis should be on thesis research. Faculty must be connected with undergraduate teaching programs, including but not exclusively those of physics departments. And, for this situation to come about, the accepted view of the role of a physics department will have to change.

The faculty in earth and planetary physics may also be involved in new undergraduate activities resulting from recent demands for greater participation by universities in the problems of society. An example concerns the interest in the question of environmental quality clearly expressed by many young people. Environmental studies in this sense cover a wide range of activities. The field is not composed simply of physics or science or engineering. Many of the most intractable problems are in the realms of economics, politics, and law. A broad undergraduate environmental major might consist of courses from a large number of departments designed to give a general knowledge of environmental and conservation problems but not to prepare people for future research in the area. We question the value of such a course of study. Rather, it should be possible to allow undergraduates to take a regular major in some discipline and, at the same time, take substantial interest in environmental studies. What is required for this purpose might be described, perhaps, as an environmental minor. However, it could be more than that. The Panel suggests an arrangement in which undergraduates taking a regular major can spend most of their remaining time in environmental studies. Environmental studies cover such a broad range that, properly handled, they can constitute a general education. An undergraduate should be allowed to acquire a general education either by studying the general education subjects distributed in the usual way or by going to a group of courses in environmental studies, ranging from physics and engineering to politics and law.

4.5 POSTDOCTORAL OPPORTUNITIES FOR RESEARCH AND STUDY

Postdoctoral programs give mature students a much-needed opportunity for quiet reflection and research during their most creative years. In a discipline such as earth and space sciences, with several interlocking fields of activity, postdoctoral research permits an individual whose graduate education has been primarily in one of these fields to extend his experience into others. For example, a student whose doctoral thesis dealt with the dynamics of oceans might well have an important contribution to make in the dynamics of atmospheres. A postdoctoral fellowship program is a desirable way to encourage this kind of interchange. Therefore, a sufficient number of postdoctoral opportunities should be available to permit the best doctoral graduates in any of the horizontal or vertical cross sections of Table IX.1 to develop their ideas in the broadest possible context in environmental physics.

There also should be postdoctoral fellowships that permit a student who has taken a PhD degree in one of the subfields conventionally found in physics departments to move into earth and space sciences. For example, a student who has recently completed his doctorate in solid-state physics might welcome the opportunity to explore the possibility of work in solid-earth geophysics. These fellowships are particularly important in times such as the present, when support for the core subjects of physics is decreasing and many young PhD's are seeking other fields of endeavor. The present situation is discussed at greater length in Chapter 8.

There are certain pitfalls to avoid in postdoctoral fellowship programs. In broad disciplines such as earth and space sciences, it is all too easy to overload a graduate program with course work, thereby frustrating the basic objective of graduate education. The widespread availability of post-doctoral fellowships could foster the idea that doctoral candidates can scarcely be expected to do any serious research and that research should be postponed until the postdoctoral stage. This form of educational inflation should be avoided.

4.6 RECOMMENDATIONS

The Panel offers to the Physics Survey Committee the following recommendations directed toward strengthening and improving education in earth and planetary physics:

We recommend that universities examine the academic structure of their graduate schools with a view to providing adequate recognition of earth and planetary physics and arranging for graduate study and research in as many cross sections of Table IX.1 as possible.

We recommend that universities examine the academic structure of their undergraduate schools with a view to increasing the emphasis on the teaching of undergraduate physics, especially by involving all faculty engaged in fundamental research in physical science, both within and outside the physics department, in the development of the undergraduate physics curriculum and the teaching of undergraduate physics. In particular, the graduate faculty in earth and planetary physics should play a much greater part in developing the physics curriculum and in teaching undergraduate physics. Through problems and examples in undergraduate physics courses, the graduate faculty in earth and planetary physics should foster awareness of opportunities for research in earth and planetary physics among undergraduate students in the physical sciences.

We recommend that universities, private foundations, and the federal government foster arrangements whereby physics professors and recent recipients of PhD degrees in physics are able to spend a summer, a term, or a year devoted to earth and planetary physics.

5 National Programs, Facilities, and Research Centers

5.1 THE NEED FOR COOPERATIVE ENTERPRISES

Although the subjects included in earth and planetary physics vary widely in character, they have in common an emphasis on the observation of conditions as they exist and change as a result of natural causes. The dimensions of the problems often require synoptic observations, which, in turn, necessitate cooperative, standardized measurements. Cooperative efforts can be organized in an informal way among observers, in a federally sponsored national program directed toward a large-scale problem and involving many scientists for a limited period, in a national facility centered around special, major instruments, or in a research center dealing with several large problems that require the continuing efforts of many scientists, usually from several disciplines.

Programs, facilities, and centers have been established to increase the level of activity, improve the quality of work, or provide instruments when the expense and staff requirements are too great for a single institution to undertake. In many subfields of earth and planetary physics, there are national centers or facilities (see Appendix IX.A). These centers pose the difficult problem of the proper balance in the allocation of limited resources between them and the single institutions with individual investigators. In practice, national centers tend to achieve a preferred position with the sponsor because of their high political visibility and because the agency staff tends to take a proprietary interest in their success. This situation has been a source of concern in the scientific community, especially during times of budgetary stringencies, when the level of activity at a center tends to compete with the support of graduate education and research.

Often it is easier to excite interest in and to obtain support for a specific cooperative enterprise than a university research effort. Therefore, the cooperative enterprise is a means of drawing attention to the needs of a subject and increasing the level of support for it—always with the hope that other aspects of and activities in the subject also will benefit.

5.2 PROGRAMS, FACILITIES, AND CENTERS

Programs, facilities, and centers have in common a concern with problems whose scale is so large that a single institution cannot act effectively and share generally in fulfilling the following needs: (a) increasing the level of activity, (b) improving the quality of work, and (c) providing expensive equipment. They differ in the way the effort is organized and managed, in the duration of the enterprise, and in emphasis. Table IX.4 illustrates these differences.

These enterprises also differ substantially in staffing. Programs are likely to have a small staff, borrowed from agencies or institutions for the duration of the program, that provides the scientific leadership, coordination, and logistic support. The staff of a facility includes a resident scientific group to aid the visitors who use a major fraction of the instrument time and to maintain and develop the capabilities of the station. A research center employs a mixture of permanent and visiting staff of high quality, with a substantial fraction of the effort directed into major problem areas selected by the center's management.

The enterprises interact with the scientific community through their advisory committees and user committees and through evaluation panels established by the sponsors. The degree of interaction varies widely, depending on the interest of the community, the aggressiveness and effectiveness of the management, the sponsor's view of his role, his capacity to evaluate or to produce evaluations, and his need for ammunition in the "battle of the budget."

TABLE IX.4 Characteristics of Programs, Facilities, and Centers

Enterprise	Organization/ Management	Duration	Emphasis
Program	Federal agency	Limited, well-defined period	Increasing the level of activity
Facility	One or more institutions	Useful life of the instrument (years)	Providing expensive equipment
Center	Many cooperating institutions	Indefinite	Improving the quality of work and increasing the level of activity

An enterprise is initiated when one or more individuals recognize scientific opportunities and national needs and devote their energies generating enthusiasm in the scientific community, arranging an appropriate organizational structure, marshalling the support of a sponsor, maintaining the interest of scientific and financial supporters, and finding and securing capable leadership for the project. The initiation process requires a few dedicated individuals, willing and able to overcome the inertia in the sponsoring system, and a need that can be presented convincingly to the scientific community and to the budget developers.

5.3 EXISTING AND PROPOSED ACTIVITIES IN THE UNITED STATES

This section deals with specific examples of programs, facilities, and centers to illustrate the way they function. These examples should not be regarded as typical, for these styles of operation vary considerably with the strength of leadership and degree of participation of the sponsors.

The Barbados Oceanographic and Meteorological Experiment (BOMEX) was part of the Global Atmospheric Research Program (GARP). It was a sea–air interaction program designed for intensive observation of characteristic properties over a 500-km square during four periods of two to three weeks each in 1969. BOMEX was a joint project of seven U.S. departments and agencies, staffed by members of government agencies, universities, and private companies, with headquarters at NOAA's Office of World Weather Systems. Although smaller than many of the space projects, BOMEX was unprecedented in complexity in atmospheric sciences and involved 1500 participants, 12 ships, and 28 airplanes.

The plans for BOMEX were developed largely by B. Davidson of NOAA following a recommendation of a panel of the National Academy of Sciences and were based on an earlier expedition in the Indian Ocean.

Although the experiment was conducted in 1969, its success is still being evaluated. It is clear that the logistics of such a complex series of observations can be managed; that extensive valuable data can be obtained economically ($20 million to $25 million); that the quality, scale, and resolution of the data have surpassed those of previous efforts and will set a pattern for the future; and that significant scientific results have been produced and many more will result from further analysis of the data.

The radio facility at Arecibo (see Figure IX.10) came into being with the support of the Advanced Research Projects Agency (ARPA) in response to a need for an especially sensitive instrument to observe the

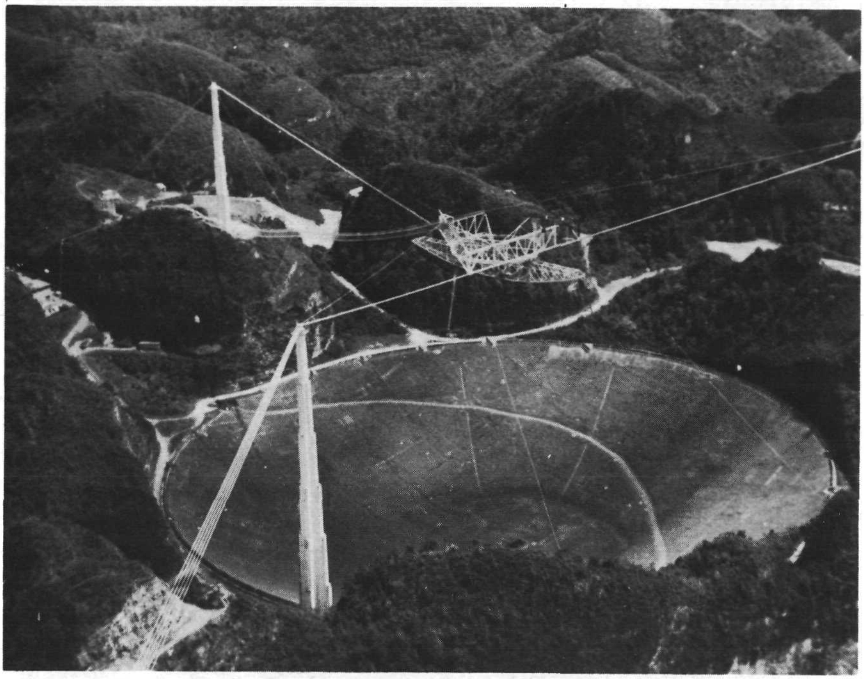


FIGURE IX.10 The 1000-ft radio telescope and Thompson scatter facility at Arecibo, Puerto Rico. [Photograph courtesy of National Astronomy and Ionosphere Center.]

upper atmosphere. The idea for the facility came from an individual; the design for it was produced by a team of Cornell University engineers supported by private companies; and the scientific support was developed at Cornell but has now spread across this country and throughout the world. The facility was placed in the tropics so that its capabilities could also be exploited in astronomy. It was built and is operated by a single institution, Cornell University. The new sponsor, the National Science Foundation (NSF) is developing it as a national center, with visiting committees to assure its continued availability to the scientific community. The scientific advances resulting from research at this facility are in the physics of the upper atmosphere, lunar and planetary mapping and ranging, and radio astronomy, particularly pulsars.

The largest and most important center in the earth sciences, now celebrating its tenth anniversary, is the National Center for Atmospheric Research (NCAR), sponsored by the NSF (see Figure IX.11). The NCAR was established in response to a clear need, recognized and described by a

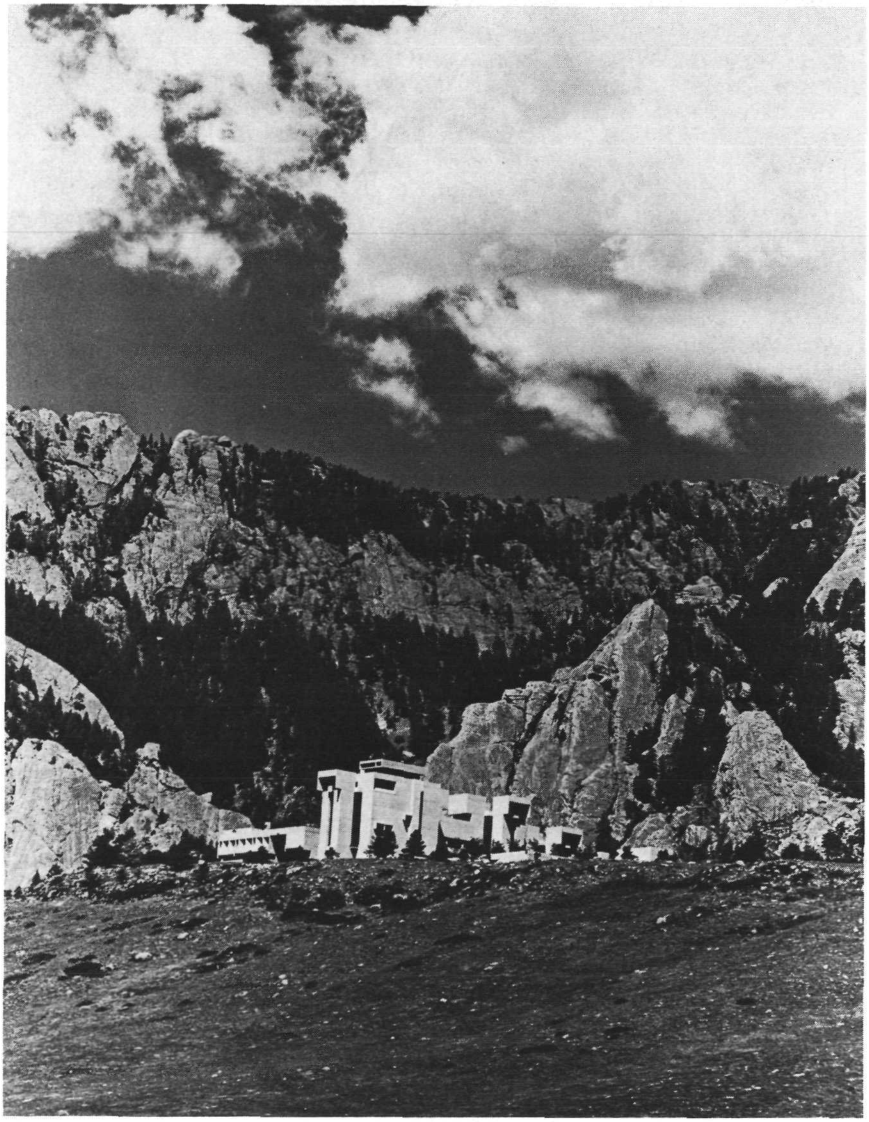


FIGURE IX.11 The National Center for Atmospheric Research at Boulder, Colorado. [Photograph courtesy of NCAR.]

distinguished group of scientists, to raise the level and quality of activity in the atmospheric sciences. It is managed by a consortium of some 30 universities under strong leadership. The goals and the evaluation of the operation have been established by a continuing series of internal committees reporting to the consortium.

Because of NCAR's maturity and size, it is important to determine how successful this Center has been. (Similar evaluations of other centers and programs also are required to ensure quality and effectiveness.) This Panel believes that such an evaluation should be undertaken by a group not directly connected with the management of the Center. This group should examine the quality of the personnel, nature of tenure review, quality of the research output, dollar value of the facilities as opposed to alternative uses of the funds, extent to which large, integrated research programs have succeeded in NCAR, and desirable future directions for the Center.

We note that NCAR has been extremely successful in building up its physical facilities, plant, and management capability. It is now a large and highly valued part of the NSF effort. It has a large research staff and able scientists in top positions.

Despite their more obvious successes, the existence of NCAR, and of centers and facilities in general, cannot be accepted as beneficial to the earth sciences without careful assessment and evaluation of their contributions. Certain invisible features of federal support suggest that a deliberate balance has to be maintained between protected, single-source support of large projects and more flexible, small-scale support of (principally) university programs. In times of retrenchment, there is a strong tendency to maintain the large projects, even at the expense of the smaller programs; when funds are more readily available, it is easier to expand the large projects, in the hope that they will stimulate and assist the smaller programs. If not carefully controlled, the process could lead to a dangerous imbalance in funding of fundamental research programs, with overconcentration on the larger efforts.

The NSF's program in deep-sea drilling, the Joint Oceanographic Institute for Deep Earth Sampling (JOIDES), is a prototype of the kind of concerted, cooperative, national program needed in the geological and marine sciences. (See Figure IX.12.) Directors of a number of major institutes cooperate in drilling the ocean floor for scientific purposes. Although the program has a single sponsor and an administrative center (Scripps), it is genuinely national and cooperative. The principal objective was to avoid the creation of new special laboratories by having the member institutions undertake various aspects of the program on behalf of the entire community.

The program for detection of underground explosions introduced a

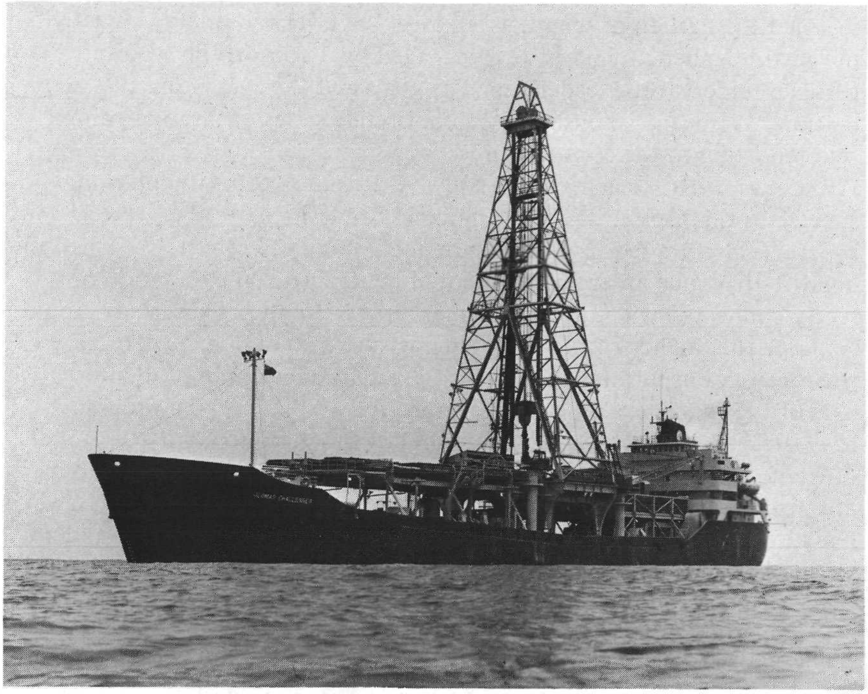


FIGURE IX.12 Deep Sea Explorer: A port-side view of the deep sea drilling project drilling vessel, *Glomar Challenger,* which is drilling and coring for ocean sediment in all the oceans of the world. The *Glomar Challenger* displaces 10,400 tons, is 400 ft long, and its drilling derrick, with a million-pound hook-load capacity, stands 194 ft above the waterline. A re-entry capability was established on June 14, 1970, which makes possible the changing of drill bits and re-entry of the same bore-hole in the deep ocean. [Photograph courtesy of the Scripps Institution of Oceanography.]

number of ARPA-financed seismic networks and data centers (see Appendix IX.B for information on these and other data centers). Reduction of support for such research threatens both the flow of new data and the future availability of existing data. Some centralized civilian program for the collection and distribution of data from worldwide networks is an obvious need in the next decade.

Fundamental studies in field geophysics tend to be limited by the cost of the field equipment, instruments, and logistic support. Typically, instrumentation in this field lags years behind the state of the art. This situation developed because most laboratories depended on their own resources for design, development, production, and deployment. A national center could well be the best way to provide expert support for the design and development of field instruments.

In the study of the region of transition from the atmosphere to space, a committee of the National Academy of Sciences has made a strong case for a major new observatory on or near the U.S.–Canadian border, with the participation of the two countries. The location provides a unique opportunity to observe those portions of the atmosphere subject to control by solar radiation and solar particles and would supplement existing observatories at Jicamarca, Arecibo, and Nancay.

For those problems requiring *in situ* observations of the upper atmosphere and nearby space, NASA has been providing the launch facilities and vehicles for scientific payloads on satellites at Cape Kennedy and on rockets at Wallops Island. Until recently, NASA, in cooperation with the Canadians, also provided rocket facilities at Fort Churchill. The planning, preparation of experiments, and data reduction from such payloads have involved a wide segment of the scientific community. Although the services provided are expensive, the organization of the efforts generally has been highly successful.

The most complicated and difficult unmanned planetary mission yet undertaken by the U.S. space program will be Viking, which will be launched in 1975 and, using orbiters and soft landers, will seek answers to the question of life on Mars. The mission has been discussed by a number of discipline-oriented committees at NASA and the National Academy of Sciences. The financial commitment, $850 million, substantially greater than earlier estimates, was secured by NASA, the Office of Management and Budget, and Congress.

The management of Viking was assigned by NASA to its Langley Center, with support from the Jet Propulsion Laboratory and other NASA centers. Industrial resources are channeled through these centers.

The scientific objectives were discussed by a NASA advisory board and led to the formation of experimental teams on various aspects of the mission. The definition of the experiments led to invitations for proposals from the scientific community and to a scientific package. In 1969, decisions were made about the experiments to be conducted in 1975.

In spite of the long lead times, the great and rising costs, and the complexity of the logistics, there are those who believe that this may be the most significant program yet proposed in the exploration of space.

5.4 IMPLICATIONS AND TRENDS IN THE 1960's

In the 1960's, expenditures increased in almost every budget category of the earth sciences, including large national programs, facilities, and centers. Difficult choices among varied activities were not required because competent work generally could be supported wherever it occurred and the scientific community was satisfied with the result.

During this period, there was a tendency to emphasize large, highly visible projects; virtually every facility or program listed in Appendix IX.A was created then. Convincing arguments suggest that the increases in support caused by the large projects brought into effect new, higher budget levels, from which the small but equally valuable individual programs profited. It clearly is easier to obtain support for large projects and to maintain them. They have definable objectives that appeal to legislators and the general public, and their budgets, once established, are less vulnerable than year-by-year grants or contracts. Possibly, however, the experience in the 1970's will be different.

This Panel does not anticipate that the importance of large projects will decrease in this decade. In every case of which we are aware, the scientific justification for the project is sound and it contributes to one or all of the values mentioned earlier in this chapter—that is, level and quality of activity and shared use of facilities. Once a facility or center has been established, it assumes its own identity. If it has a strong director and an active board, it becomes a major power in the development of scientific policy for the relevant discipline. The most likely trend in the 1970's is not a decrease in the number of facilities and centers or in their level of support but increasing difficulty in establishing new enterprises.

In regard to two specific topics, the Panel has recommendations for future procedures; these are the balance of activities in NSF programs and the coexistence of large and small science projects within a single budget of NASA.

The NSF is the most important single source of support for academic research and is also strongly committed to the concept of national centers. Besides NCAR and Arecibo in the atmospheric sciences, it supports Kitt Peak National Observatory and the National Radio Astronomy Observatory in astronomy (Figure IX.13). All have been successful, and the pressure for new centers or the expansion of existing centers continues. Current proposals for new centers include the new U.S.–Canadian aeronomy facility and the less-well-defined concept of a facility for geophysics.

National centers have expensive plants to maintain, tenured and nontenured staff, and continuing research programs, all of which involve a long-term commitment of support from the NSF. Since the NSF is generally their sole source of funds, it has a direct responsibility that has no parallel in university research and other activities supported under NSF discipline-oriented programs. In these other activities, the NSF acts as a partner, often providing most of the funds but without direct responsibility for tenure or the long-term health of the programs. The important

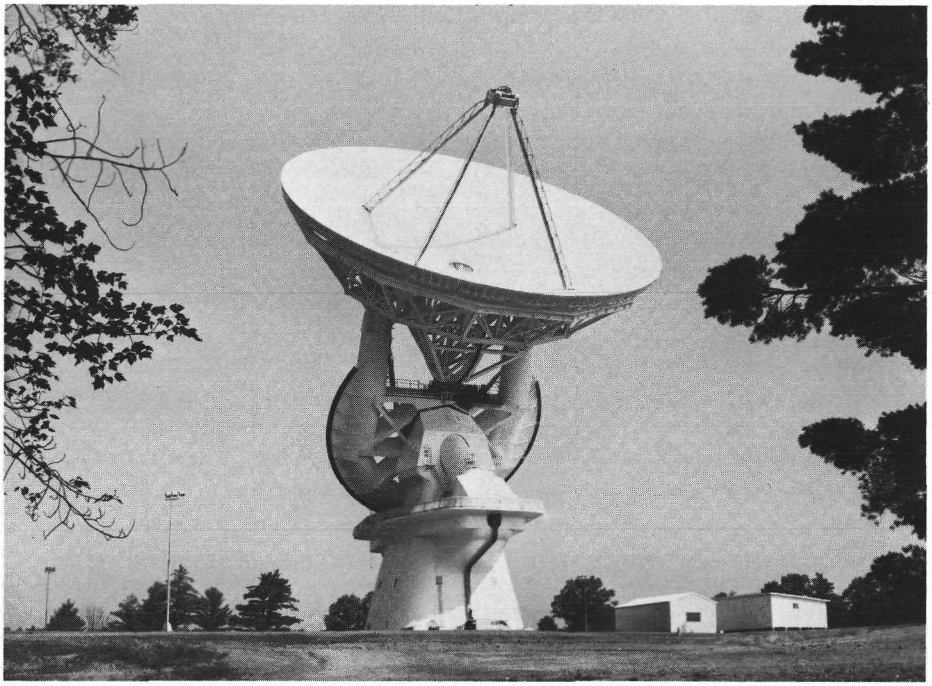


FIGURE IX.13 The 140-ft antenna at the National Radio Astronomy Observatory, Green Bank, West Virginia, the largest equatorially mounted radio telescope in the world, has been operated successfully down to a wavelength of 1 cm. [Courtsey of Associated Universities, Inc.]

problem that could develop in the 1970's concerns the possible effects of supporting two programs of such different character during periods when agency budget allocations are lower than in the 1960's or decreasing.

When agencies have a dual responsibility to in-house and extramural research, the predictable response to budgetary restrictions is to withdraw first from the extramural activities. Nevertheless, it may not be the best response in terms of the overall national science capability. The effect of current withdrawal of support for the universities by the Department of Defense, the Atomic Energy Commission, and other agencies is serious, and it would be more serious but for the steady and now increasing support provided by the NSF. If budgetary constraints were to force the NSF to react as other agencies have, irreparable harm could result. Therefore, we believe that caution must be exercised in the expansion of existing programs or the creation of new programs for which the NSF has direct managerial responsibility. A balance must be sought that makes some allowance for the future possibility of fiscal stringencies.

It cannot be assumed that, faced with fiscal restrictions, the NSF could act differently from other agencies. A mild precursor is the expenditure ceiling imposed on the universities since fiscal years 1969 and 1970. This example is complicated to analyze because it did not result in an overall reduction in university expenditure. However, in specific cases, it did result, without warning, in retroactive cuts amounting to as much as 40 percent and thus to substantial decreases in research activity. Possibly some relief could have been afforded to atmospheric sciences programs by decreases in in-house activity in NCAR. However, as Chapter 8 indicates, the operation and program equipment budget of NCAR rose during this period (from $9.8 million in 1968 to $10.4 million in 1969 and to $11.0 million in 1970).

In the creation of new centers, the NSF acts mainly in response to the expressed wishes of the science community, generally voiced through the NAS. *We recommend to the Physics Survey Committee that the National Academy of Sciences devote increased attention to the review of panel, committee, and board recommendations for new centers, programs, and facilities, particularly to the long-term implications of these proposals with respect to the balance of science activities supported by the National Science Foundation.*

Effective reviews concerning national centers, whether undertaken in the NAS or the NSF, require objective external assessments of the success with which centers have been able to achieve their objectives. Such information does not exist, at least not in available or satisfactory form. *We therefore recommend that the Physics Survey Committee transmit to the National Science Foundation our recommendation for a continued external review of the work of its centers.* This review should have the same depth and quality as that used for individual research grants, and the procedure should aim at maximum comparability.

The NASA Space Science and Applications program contains a mixture of very large projects, the character and time schedule of which cannot be easily modified (for example, Mars Viking, Outer Planets, High Energy Astronomical Observatory), smaller projects that, although costly, are relatively flexible (for example, Planetary Explorers, Small Astronomical Satellites, Pioneers, and small earth satellites), and low-cost items (laboratory and theoretical research, rockets, balloons, and aircraft). Problems arising from this mixture of different types of programs within a single budget have some analogy to those encountered with national centers. The large project is much easier to maintain, and, once started, the budgetary stability provided is important for all involved in the NASA programs. The large project is also easier to justify to a nonscientific audience, for it often will have clear-cut objectives and will offer major technological challenges. The NASA managers, therefore, tend to look favorably on large missions and projects.

Large missions require definite sums of money on a definite time schedule and they cannot be readily delayed or decreased in size. Thus, in the event of fiscal problems, there is a strong tendency to reduce other parts of the program first; this policy will often mean cuts, delays, or elimination of the flexible programs of lower cost.

There is no single opinion in the scientific community as to the relative merits of such small and large programs as the Large Space Telescope and the Small Astronomical Satellite, or Viking and Planetary Explorer. However, successive NAS studies, which are discussed in Chapter 8, have emphasized the value of the smaller programs and have insisted that they must not be sacrificed to support more expensive missions. The most recent study emphasizing this point was *Priorities for Space Research 1971–1980,*[11] which placed Planetary Explorer and increased rocket and balloon astronomy before Grand Tour. NASA's response in the 1972 budget was to request Grand Tour but not the two lower-cost items. Grand Tour, however, was subsequently replaced by a less expensive program.

In this Panel's opinion, the secure future of the space science program demands a close and effective working partnership between NASA and the scientific community. The conflict between managerial and scientific requirements in the initiation of new missions could prevent such a development. *We therefore recommend to the Physics Survey Committee that the Space Programs Advisory Committee of the National Aeronautics and Space Administration and the Space Science Board of the National Academy of Sciences initiate and maintain a discussion with the National Aeronautics and Space Administration management on the problems and priorities of small and large missions, with the aim of achieving a common policy in this important area.*

6 The Role of Computers

6.1 GENERAL CONSIDERATIONS

The development of high-speed, large-storage-capacity computers has benefited all physics. It has made possible the solution of complex equations and thus the quantitative description of more realistic models of physical systems. In addition, it has permitted on-line analysis of experimental data and the automatic and remote programming of experiments in accordance with the analysis in real time of the results of an experiment. However, in geophysics, planetary physics, and space physics, the computer has an even more important role to play. The environments described usually involve many variables and interrelated equations. To determine the behavior of the system, it is necessary to acquire and organize large amounts of data; physical models of the systems also are exceedingly complex. In this chapter, we are particularly concerned with these special aspects of the role of the computer.

Knowledge of the interior of the earth has been obtained largely from seismological observations. The prime tools are the earthquake-generated elastic waves that propagate through the interior and as surface waves. Networks of high-frequency instruments (about one cycle per second) make it possible to determine travel times of various types of geometrical rays to within one or two seconds. Instruments of intermediate period [20–200 cycles per hour (c/h)] are used to record the wave groups of surface waves, from which one determines the dispersion of the surface waveguide. From broadband instruments (0.5-60 c/h), it is possible to resolve the frequencies of many of the free oscillations of the earth when excited by very large earthquakes. In addition, the mass and moment of inertia of the earth are ascertained from astronomical measurements and provide additional constraints on the possible internal structures of the earth.

For each of the types of seismic data, a theoretical formulation exists that permits the calculation of synthetic results if the earth model is specified. Since the earth model is nearly always considered to be spherically symmetric, or slightly perturbed from spherical symmetry, these calculations require the solution of systems of ordinary differential equations by numerical integration in the radial variable. Generation of synthetic data for one earth model—about 500 synthetic measurements—takes about 45 minutes on a CDC 3600. Determination of an earth model that matches actual data requires the use of a converging iteration in model parameter space and can be made quickly if the starting model is linearly close to the true one. Computation of the resolving power and uniqueness of an earth model, as determined by a set of data, requires further calculations of an eigenfunction, involving the inversion of large matrices; the procedure can consume more than an hour of computer time. The extension of these methods to cover structural perturbations in the geographical coordinates involves an increase of a factor of 10 in computing requirements for even the more restrictive models. Such investigations are anticipated in the next few years.

Another task for which computers are heavily used in seismology is the reduction of data. There are now several two-dimensional receiver arrays, designed to operate with the same directional capability as a comparable microwave antenna, to resolve the directional spectrum of incoming signals. In the largest of these, 625 sensors are digitally recorded, and the data reduction system must synthesize various beam outputs both in real time and later. The requirements of this work challenge both the speed and the mass-storage capabilities of modern computing machines. In addition, the present lack of suitable interactive graphic facilities has hampered practical access to large data libraries.

The atmosphere and the oceans form a coupled system of geophysical fluids that presents an enormous challenge to the computer. The economic benefits that would result from the reliable prediction of the behavior of the atmosphere for as much as two weeks in advance would be great. In a typical massive program designed to simulate the atmosphere, grid points some 200 to 300 km apart in the atmosphere can be selected. At every grid point, the model will include the properties at several heights in the atmosphere. The basic problem, given an initial set of conditions at every grid point, is to determine the future behavior of the atmosphere by integrating the equations of fluid dynamics, beginning with these initial conditions. The accuracy of these integrations necessarily becomes degraded with time, since the atmosphere contains a considerable amount of energy in turbulent length scales that are comparable to and smaller than the grid spacing. Present estimates are that a major degradation will occur in the prediction after a characteristic time interval. (The next section of this chapter describes the progress and prospects in the development of numerical models of the atmosphere.)

The potential of the computer for handling the large amount of computation needed in numerical prediction stimulated development of the Global Atmospheric Research Program (GARP). One objective of this program is to measure weather conditions at the necessary grid points, thus providing sufficient initial data for a realistic numerical prediction. This research program will make use of observations from satellites and various ground and atmospheric sensors and will provide a continuing flow of information that must be fed into the calculation of the behavior of the atmosphere as a series of corrections. The GARP research program eventually will require the use of computers having speeds two orders of magnitude greater than those of the highest-speed commercial computers currently available. The design and construction of such computers are under way.

In several areas in atmospheric physics, understanding of phenomena has not yet approached the level that will be required for reliable prediction of the future state of the atmosphere. The scientific problems that necessitate further study include energy and momentum transfer in the lower boundary layer of the atmosphere (the lowest 1000 m), cloud convection, clear-air turbulence, tropical cloud clusters, internal boundaries or "fronts," and various kinds of gravity waves related to them. Thus computers play an essential role in meteorology; they are needed not only for weather prediction but for creation of a worldwide system of observations in which data can be fed continuously into them. It is hoped that a better understanding of some of the fundamental mechanisms in the atmosphere, for which the physics has not yet been clarified, will result.

The air–ocean interface is still little understood. Conditions at this interface play an important role in the global atmospheric prediction problem through the transfers of sensible heat and water vapor between ocean and atmosphere and through shear stresses and wave generation. Large amounts of the incoming solar energy are transported by ocean currents; therefore, these oceanic processes are coupled to the behavior of the atmosphere and must be understood at the same time that the behavior of the atmosphere is being calculated. Computers also will be important in future research in oceanography, since the ocean is a fluid that is subject to a wide variety of motions, including turbulent motions on various scales. Here, too, there is a major problem of prediction. Scientists are now investigating much of the physics of the ocean with the help of computers. Numerical techniques, similar to those used in meteorology, are important in understanding the response time of the ocean to varying energy inputs. Spectral analysis of long-time series of waves, currents, turbulence, and the like is beginning to yield clues to the energy distribution and the transfer of energy up and down the space scales of motion.

The role of the computer is crucial in data reduction for space experiments. Spacecraft investigate a wide variety of phenomena—the magnetosphere, interplanetary plasma, solar radiations, cosmic rays, earth, and stars. Satellites and space probes generate a continuous flow of information that must be received on the ground and stored on magnetic tapes. Almost without exception, experimenters with packages on board spacecraft have been inadequately prepared to handle the large flow of data telemetered back to them. They have given insufficient attention to writing the computer programs necessary to use the data in the most efficient way and to extract the relevant physical information from them. As a result, data processing for most of the information obtained from spacecraft often is delayed for two or three years. This situation is not entirely the fault of the experimenters, for the necessary computer time frequently is limited, unavailable, or prohibitively expensive.

The pictures of the surface of Mars returned from the Mariner VI, VII, and VIII spacecraft provide examples of extremely complex data handling. These pictures are coded, picture element by picture element, in compact form. The brightness of each picture element affects the brightness of neighboring elements, since, to obtain the greatest return of scientific data, the engineering system was designed to minimize the large-scale contrast of the surface. Therefore, reconstruction of these pictures to a normal appearance presented a major problem in data processing. Final pictures of the surface of Mars are not ready until about one year after the acquisition of the data. While this delay might seem unfortunate, one should consider that the data stream from the Mariner spacecraft contains a great deal more scientific information than would have been available had the pictures been sent in a more conventional television mode.

In the case of many spacecraft measurements on particles and fields, much preliminary data handling is now done by computer, on line, before the raw data are delivered to the experimenters. These preliminary steps can include the preparation of computer-generated graphs of the data or, in more sophisticated cases, the generation of motion pictures or color photographs.

Field stations frequently are not conveniently located. Even when operated by a university, a field station may be far from the campus. Students and staff formerly lost much time in commuting to field stations for experimental runs. Now such trips are largely unnecessary. By means of computers, remote terminals, and land lines (or microwave links), most work that one might wish to do at a field station can be accomplished directly from the campus, including rotation of antennas, adjustment of instruments, and recording and processing of observations. At the University of California, San Diego, a solar-wind observatory has been set up in this way. Antennas and auxiliary equipment as much as 100 km from the campus can be directly controlled from the laboratory. An hour of observations can be conducted with no more difficulty than that involved in going to a seminar or the library. Of course, it is necessary to have technicians visit the site occasionally; responsible scientists also must be present when new experimental equipment is installed. However, much of the travel to and from field stations for purely scientific reasons has been eliminated. This technique of remote control of field stations could well be extended; national observatories can and should be so organized that experiments at these facilities can be conducted from almost any campus with the relevant equipment and interests.

6.2 NUMERICAL MODELS OF THE ATMOSPHERE

Fundamentally, the atmosphere is a turbulent fluid, and its study during the last two decades is an unusually clear example of a revolutionary change produced by the development of computers. It has no simple regularities that would make possible the prediction of its motion as the tides are predicted. Richardson, in his pioneering book on numerical weather prediction,[3] remarked that "the Nautical Almanac, that marvel of accurate forecasting, is not based on the principle that astronomical history repeats itself in the aggregate. . . . Why then," he asked, "should we expect a present weather map to be exactly represented in the catalogue of past weather?" He proposed instead to predict the weather by integrating the nonlinear partial differential equations governing atmospheric motion numerically, just as the Almanac is produced by solving the ordinary differential equations of celestial mechanics. After estimating that 64,000 human computers equipped with slide rules and desk calculators would be needed to trace the weather for the entire globe, he indulged in a fantasy in which he visualized this army of computers as an orchestra conducted by a master computer in a hall surrounded by "playing fields, houses, mountains and lakes, for it was thought that those who compute the weather should breathe of it freely."

For reasons not then apparent, Richardson's hand calculations were wide of the mark, and, because his proposal was considered quixotic, it lay dormant for 24 years. His book had gone out of print when, in 1946, John von Neumann renewed the attempt to integrate numerically the gas dynamical equations for the atmosphere, this time with the aid of a high-speed electronic computing machine and the additional empirical and theoretical knowledge of atmospheric mechanics developed during the past 25 years. The work of C. G. Rossby was useful in this effort.

Von Neumann assembled a group of theoretical meteorologists at the Institute for Advanced Study at Princeton University under the leadership of J. Charney, who had devised a method of overcoming the numerical difficulties that led to the errors in Richardson's effort. A compressible, stratified fluid, held gravitationally to a rotating earth, can support a variety of motions, including acoustic- and inertial-gravity waves, which are of little meteorological importance but which greatly complicate the solution of the gas dynamical equations. Charney filtered out these unwanted "noise" motions by mathematically incorporating the principle that slow, large-scale motions in a rotating system are close to equilibrium in the sense that the accelerations due to pressure, gravity, and Coriolis forces dominate purely inertial effects. A hierarchy of mathematical models of increasing complexity was constructed to embody more of the physical properties of the atmosphere. The first integrations were performed on the ENIAC* in

*Electronic Numerical Integrater and Computer, the first fully electronic computer, designed primarily for ballistic calculations by Eckert and Mauchly at the University of Pennsylvania and later equipped, under the guidance of von Neumann, with a logical control that converted it to a general-purpose computer.

1950, with what was essentially Rossby's model, in which the flow of the atmosphere at a midlevel is idealized as two-dimensional and incompressible. These calculations sufficiently resembled real conditions to encourage further experimentation. After the completion, in 1952 at the Institute for Advanced Study, of a new electronic computer that had been designed by von Neumann and J. Bigelow, it was possible to model the vertical structure of the atmosphere and to deal with potential-kinetic energy conversion as well as the horizontal shear instability and dissipative properties of the atmosphere. In 1953, the first prediction of cyclogenesis, regarded then as the principal problem of meteorology, was obtained with a model that permitted conversion of potential to kinetic energy but contained no sources or sinks of energy. This result had two consequences: it tended to confirm Charney's earlier explanation of cyclogenesis as due to horizontal temperature gradients in midatmosphere, and it interested the U.S. Government in the possibilities of an operational numerical prediction system. A numerical weather prediction unit was established in 1954 and began operations in 1955. Since then, many other countries have established similar units.

Following the numerical prediction of cyclogenesis, it was natural to study the interactions of the cyclone wave with the zonal flow. H. Jeffreys, V. Starr, and others had shown that the wave disturbances acted as turbulent eddy elements transferring angular momentum (against the gradient) to the mean zonal flow. In seeking a mechanism for this phenomenon, Charney proposed the hypothesis that the cyclone waves, while deriving their energy *from* the unstable zonal flow through release of the potential energy associated with the poleward temperature gradient, are also required by the stabilizing effect of horizontal shear to return energy in kinetic form *to* the flow and thus maintain it against dissipation by friction. This hypothesis was strikingly confirmed in a numerical experiment conducted at the Institute for Advanced Study by his associate, N. Phillips. Phillips introduced into the simplest of the three-dimensional models a heating function varying uniformly with latitude and frictional mechanism. As a result, a broad westerly current with a uniform poleward temperature gradient was generated. As predicted, this current became unstable and developed wave perturbations that transferred kinetic energy back to the current and caused it to become narrow and intense, as in the observed westerlies. For the first time, the principal dynamical elements of the general circulation of the atmosphere were combined into a single mathematical model, which opened the way for a numerical attack on the problems of long-range prediction and a dynamical theory of climate.

With increased understanding of the mechanism of interaction between inertial-gravity waves and the low-frequency, energy-bearing flow of the atmosphere, it was possible to return to Richardson's work on the primitive gas dynamical equations with new insights. After preliminary experimentation, the first simple model for this purpose was constructed at the Institute for Advanced Study, and integrations were developed under the direction of J. Smagorinsky in a new unit of the U.S. Weather Bureau in Washington, D.C. The appearance of a strange nonlinear computational instability at first prevented long-term integrations. The purely mathematical source of this instability was located by Phillips, and a finite-difference algorithm was designed by Arakawa to overcome it. Smagorinsky's group at the Weather Bureau and Mintz and Arakawa at the University of California at Los Angeles then performed the first long-term integrations for periods of a month or more. As more of the physical properties of the atmosphere and the earth's surface were incorporated—cloud, condensation, radiative heat transfer, frictional boundary layers, mountains, oceans—the outlines of observed climate began to emerge. Currently, the gross features of the earth's climate, although not its climatic variations, are understood. Bryan and Manabe of Smagorinsky's group have attempted to construct numerical models of the coupled atmosphere–ocean circulation. Although these integrations have yielded a modicum of success, many problems remain. Foremost among them is the lack of knowledge of the mechanism of turbulent heat flow in the main ocean thermocline and the nature of turbulent diffusion in the ocean in general. In addition, even though the characteristic time scales of the atmospheric motions are far less than those of the oceans, the atmosphere cannot be regarded as in statistical equilibrium with the underlying ocean surface on the time scale of significant changes in ocean circulation. The nature of such interacting systems is not understood. Their study will challenge computational capabilities and require computers of extremely high speed and capacity.

As the models grow in verisimilitude, and therefore in complexity, they impose greater and greater demands on computing facilities. That the conventional methods of obtaining data—primarily by balloon soundings of temperatures, winds, and moisture from land-based stations and weather ships—are inadequate is increasingly apparent. Less than 20 percent of the atmosphere is adequately observed. Influences propagating from uncharted regions of the atmosphere affect predictions for all points on the earth within two or three days. Fortunately, development of the meteorological satellite occurred at the time when the requirements of numerical prediction began to create the need for the increased data coverage. Spectroscopic measurement of the temperature-dependent thermal emission of CO_2 and water vapor in various spectral intervals, with varying photon mean free paths, makes possible the inversion of the equations of radiative transfer to obtain the temperature as a function of pressure, and thus of height, in the atmosphere. Although clouds of any appreciable thickness are opaque in the infrared, high-resolution infrared spectrometry permits the sensing of temperature through broken clouds, and microwave spectrometry makes possible the sensing of temperatures through all but the heaviest rain clouds. The United States currently is participating with other countries in the United Nations-sponsored GARP, a program whose principal objective is to devise a satellite-based global observational system for measuring the large-scale field of motion of the atmosphere. Present plans call for a mixed observational system, consisting of remote sensors, carried by satellites, and immediate sensors, carried by balloons and buoys but located and interrogated by satellites. Charney, Halem, and Jastrow have shown that, in principle, it is possible to infer winds and pressures from measurements of temperature and moisture alone; whether this can be done with the required accuracy is still uncertain.

Whatever the ultimate global observational system, the task of real-time reduction of the data will be enormous and will require new, highly sophisticated processing techniques as well as large, fast computers. Existing meteorological satellite information-processing requirements emphasize this need. Geosynchronous satellites permit high time resolutions; they are therefore suited to sensing the smaller-scale motion field for both detailed local forecasting and scientific investigation of mesoscale atmospheric phenomena. The quantity of information involved again is exceedingly large and will challenge computer capabilities and capacities.

The growing knowledge of the constitution and energetics of the atmospheres of the inner planets, resulting from space probes and earth-based measurements, has stimulated attempts at numerical simulation of their circulations. Loevy and Mintz have presented numerical integrations for Mars, and attempts, inspired by Goody's and Robinson's suggestion that the high surface temperatures on Venus may be produced by a deep circulation, are now under way to calculate the Venus circulation. Similar attempts undoubtedly will be made for the atmospheres of the outer planets, Jupiter, Saturn, and Uranus, as soon as their constitutions can be reasonably inferred. The photosphere of the sun is another probable candidate for numerical simulation.

Scientists continue to encounter problems at the earth's surface, all of which are associated with turbulence. What are the statistical properties of the gross turbulent structure of the atmosphere? What can be said about the natural variability of climate? To what extent is this variability an intrinsic property of an atmosphere with fixed boundary conditions and fixed constitution, and to what extent is it determined by changes in boundary conditions and constitution? The transfer of mass, energy, and momentum in the atmosphere results from anisotropic turbulence in the surface boundary layer and also in midatmosphere, especially at the base of the stratosphere in the so-called planetary jet stream. Cumulus convection (gravitational instability produced by release of latent heat in upward-moving currents of air) is an important energy-supply mechanism for the atmosphere, especially in the tropics. Each turbulent phenomenon requires separate study before it can be successfully incorporated, through parametric equation, into a mathematical model of the atmospheric circulation. Each is amenable to numerical computation and, to a first approximation, requires the same amount of computer time, independent of scale. Thus, estimates suggest that the times for the calculation of the general circulation or for a cumulus cloud complex are about equal. A major part of GARP will be devoted to the field study of small-scale atmospheric turbulence of various sorts, with emphasis on the organization of tropical convection, one of the weakest links in present understanding. The findings will contribute directly to the design of numerical models for the tropics.

The possibility of incorporating turbulent fluxes parametrically in numerical models of the large-scale circulation depends on the distinctness of the scale separation between the turbulence and the large-scale motions. Probably, the energy in the atmosphere, inserted mainly at large scales, does not flow in appreciable amounts to smaller scales by nonlinear interaction, as postulated by Kolmogorov for homogeneous, isotropic turbulence; rather, the large-scale energy is dissipated in surface or internal boundary layers. As a consequence, the energy per unit horizontal wave number decreases with wave number at a much higher rate than that predicted by the Kolmogorov theory, at least until high wave numbers are reached.

Lorenz, Kraichnan, and Leith have shown in separate studies that the degree of predictability of the atmosphere as a deterministic mechanical system depends heavily on the spectral structure of its velocity fluctuations and consequently on the mechanism of dissipation. Therefore, the rate at which uncertainties in initial conditions due to small-scale turbulence propagate toward large scales depends on the mean spectrum of the atmospheric energy. In addition, observational error influences the large-scale instabilities directly. The best estimates of the rate of growth of error have been obtained from numerical experiments with a variety of computer models. These place the limit of determinacy for the largest-scale motions at between one and two weeks. Not yet known is the extent to which predictions of averages or distributions are possible for longer times, taking into account variations of surface boundary conditions, especially of ocean temperatures and snow and ice cover. The statistical effects of modifications in the atmospheric circulation produced by small changes in boundary conditions or constitution, for example, by small changes in ocean surface temperatures or CO_2 content, cannot be ascertained until the statistical behavior of the undisturbed system is known. Such calculations will require computers of the highest speeds and greatest capacities.

Improved knowledge of small-scale atmospheric processes, especially of turbulence, together with better observations, should result in improved weather prediction for intervals up to two weeks. The extent to which seasonal variations of space and time means and distributions will remain predictable is not yet known. A suitable theory of climate and climatic change will require that the atmosphere and oceans be treated as a coupled system. The investigation of the combined ocean–atmosphere circulation has barely begun; much work lies ahead.

7 International Cooperation

Of all the sciences, those concerned with the physical environment have had the longest history of explicit and extensive international cooperation. These sciences are planetary and interplanetary in nature, linked to solar radiations, particles, and fields. Early in their history, there developed a need for synoptic data observed at many points on or above the surface of the earth over significant periods of time. There also was need for coordinated multidisciplinary observations at the time of discrete, short-lived phenomena, such as solar flares or magnetic storms, and of their varied effects on the earth's environment.

Gauss recognized the need for geomagnetic data from observatories other than his own and, with Humboldt, created the Magnetische Verein in 1836, which brought about cooperation among scientists in several European nations in making simultaneous measurements of the earth's surface field. Some 25 years later, the Mitteleuropäische Gradmessung of 1862 led to the establishment of the International Union of Geodesy and Geophysics (IUGG), now one of the 15 unions making up the International Council of Scientific Unions (ICSU).

Man's ancient concern with weather stimulated cooperation in meteorology in the last century. In 1853 a group of seafaring nations formulated an oceanic weather observation program in the interest of safety at sea. The International Meteorological Organization was established in 1873 to extend such observations to land areas, and a scheme for exchange of weather information among nations was devised. Crucial to the usefulness of such exchange, if timely forecasts were to be made, was the invention of the telegraph and later the wireless. Commercial aviation imposed new demands for weather information and provided an important stimulus to meteorological studies. Such social and economic interests in weather led, in 1951, to the formalization of cooperation by the establishment of the World Meteorological Organization (WMO) as a specialized agency of the United Nations (UN). The WMO now consists of 133 member nations and territories; it has been active in promoting worldwide networks of surface and upper-air observations, standardization of instruments and methods of observation, intercomparisons, and other procedures.

Historically, the international exchange of synoptic meteorological information has been a model of open and free access. Weather observations are transmitted and made available to member and nonmember nations alike. Data are freely exchanged between the eastern and western European countries. In Asia, meteorological information from mainland China has long been available.

Seismology provides a third historical thread of interest. As with meteorology, the motivation had its roots in catastrophes. Scholarly enquiry within the Society of Jesus led to observations of earthquakes in many parts of the world and a tradition of exchange that continues to the present. In this century, the establishment of the International Association of Seismology and Physics of the Earth's Interior—a part of the IUGG—owes much to the Jesuits in the recording of earthquake observations at scattered sites and in the exchange of data. These activities in seismology led in the 1960's to the establishment of the International Seismological Centre in Edinburgh, which prepares catalogues of earthquakes based on observations from stations all over the world. In the same decade, under U.S. auspices, a World-Wide Network of Standard Seismic Stations was established; it consists of some 115 stations well distributed around the world.

7.1 MECHANISMS FOR COOPERATION

A variety of mechanisms exists for furthering international cooperation in the earth sciences. These range from relationships established personally and informally by individual scientists or groups of scientists through more formal nongovernmental mechanisms, particularly those embraced by the International Council of Scientific Unions, to formal governmental arrangements within the UN and its specialized agencies, such as the WMO, the World Health Organization, the International Atomic Energy Agency, UNESCO, and the UN Committee on Peaceful Uses of Outer Space. Nor do the preceding institutions exhaust the available mechanisms. Many bilateral programs are undertaken—for example, those of NASA with

Japan and a number of European countries—while the National Academy of Sciences not only works with ICSU and its unions but has active ties to sister academies in many parts of the world.

Some distinction may be drawn between the governmental framework of the UN and its agencies and the complex represented by ICSU, its subject-matter-oriented unions, and various committees and commissions. The UN agencies have a political character, for their members are governments. The ICSU organizations are essentially organizations of scientists; their functions are to serve the interests of science as perceived by scientists. However, there is significant cooperation between these governmental and nongovernmental groups: for example, the UN Committee on Peaceful Uses of Outer Space turns to ICSU's Committee on Space Research (COSPAR) for advice on scientific matters; UNESCO provides subventions to ICSU and uses its counsel in many ways; both UNESCO and WMO were active in the International Geophysical Year (IGY) and successor programs developed within ICSU; and WMO and ICSU jointly are engaged in the planning of the Global Atmospheric Research Program (GARP).

Within ICSU, there are several organizations dedicated to international cooperation in the space and earth sciences, particularly (among the unions), the International Union of Geodesy and Geophysics (IUGG), the International Union of Radio Science (IURS), the International Astronomical Union (IAU), the International Union of Pure and Applied Physics (IUPAP), and the International Union of Geological Sciences (IUGS), and (among ICSU's specialized committees and commissions) the Committee on Space Research, the Scientific Committee on Antarctic Research, the Scientific Committee on Oceanic Research, the Upper Mantle Committee, the Inter-Union Commission on Geodynamics, and the Inter-Union Commission on Solar-Terrestrial Physics. Here again, an admittedly oversimplified distinction can be made between unions and committees. The primary functions of the unions are to provide (a) a forum for communication among the world's scientists within the same or closely related disciplines, through scientific sessions at general assemblies and symposia sponsored by subgroups of the unions; (b) a mechanism for planning cooperative programs; (c) a mechanism for the establishment of international standards in fundamental units, measurements, and calibration as well as in nomenclature and publication; and (d) a means for issuing some types of publications—for example, astronomical and geomagnetic indices, proceedings of symposia, and international journals (such as *Acta Crystallographica*). The special committees are specifically concerned with fields that combine two features: (a) an interdisciplinary element involving two or more unions and

(b) the actual or prospective existence of an active, declared international effort such as the IGY, the International Year of the Quiet Sun (IQSY), or the Upper Mantle Project (UMP).

7.2 THE INTERNATIONAL GEOPHYSICAL YEAR

The IGY was a remarkable collaborative international endeavor, not only because it accomplished much scientifically and engaged some thousands of individuals in a rewarding venture, but also because it gave rise to several major international programs of continuing interest. The IGY was possible because the earth sciences were prepared for a major advance and new technical capability was combined with the enthusiasm of active research scientists.

Two antecedents of the IGY demonstrated the value of cooperation among nations in geophysics and the organizational feasibility of conducting such efforts. Compared with the IGY, both were small and restricted in scope. In the First International Polar Year of 1882–1883, a dozen nations worked together almost solely in the high northern latitudes to study surface meteorology, geomagnetism, and auroral physics. Perhaps the most significant result of this modest endeavor was Fritz's charts of the aurora, which delineated the auroral configuration around the north geomagnetic pole. The experience of 1882–1883 led to the Second Polar Year of 1932–1933, in which the relatively new field of ionospheric physics was added to the scientific agenda. A score of nations worked together, again with considerable emphasis on the high northern latitudes but with a sprinkling of stations at middle and southern latitudes. The promise of this venture was not fully realized because it was plagued by the depression of the 1930's and had almost negligible support, but the ongoing activities of such organizations as weather bureaus in various countries, considerably stronger than in the 1880's, permitted the conduct of significant investigations. Again, if a single example can suggest the character of the results, such an example might be the linking of disturbances in the ionospheric layers to geomagnetic and auroral activity. Only a few years before, Breit and Tuve had invented the radar technique of probing the ionosphere by pulsed emissions of radio energy. In the Second Polar Year, use of this technique provided insight into the nature of the ionospheric layers and thus into problems in radiowave propagation and communications.

The background of the two polar years—there had been proposals that similar ventures might be fruitful at 50-year intervals—was well known to the initiators of the IGY, who felt that the time was ripe 25 years after the Second Polar Year to conduct a far more critical and intensive study in the planetary and space sciences. The rationale behind this conviction rested on a number of historical and contemporary events, of which the following appeared most relevant. In many of the fields, decades and even centuries of observations had led to some understanding of at least gross features and, in some disciplines, to useful hypotheses and theories. A worldwide gathering of synoptic data was feasible because of the capabilities of several nations; in short, the collection of data on a global basis seemed feasible in terms of manpower and costs, given the cooperation of 40 to 60 nations. The development of high-speed computers made it possible to handle the anticipated data. Electronics, relatively new in the 1930's, had developed enormously during the war years, providing sensitive devices, small and reliable, for measuring with precision a wide variety of parameters, while such new tools as rockets seemed certain to open new fields. It appeared probable that creative investigators would commit themselves to the prospective international effort. In fact, leading geophysicists did enter into the IGY program from its planning period (1951–1957), through its operational interval (1957–1958, and its one-year extension through 1959), to the years following, when data were analyzed. The scientific rewards, as perceived by these active scientists, including younger ones who became attracted to the IGY effort, then led to the development of successor programs in more specialized areas (for example, the IQSY and the UMP).

The IGY embraced 13 areas of activity: aurorae and airglow, ionospheric physics, geomagnetism, cosmic rays, solar activity, meteorology, oceanography, glaciology, gravity, longitude and latitude determinations, and seismology; rockets and satellite programs concerned with the upper atmosphere and near space were added as special areas of activity. The emphasis was first on those fields now associated with the term solar-terrestrial physics (disciplines concerned with the upper atmosphere, the medium beyond, the driving solar forces); second, on interface phenomena associated with the earth's heat and water store; and, third, on solid-earth geophysics.

Thus, seismology and gravity entered into the program largely because IGY expeditions would go to regions not ordinarily frequented by man (especially Antarctica), and these opportunities should be made available to seismologists and geodesists, whose interests for somewhat differing reasons did not, strictly speaking, demand the IGY type of measurements and experiments in concert. Nonetheless, this exposure to the IGY approach was to bear fruit in both IGY results and the subsequent UMP, which stimulated the present cooperative program on the dynamics of the solid earth.

7.3 SOLID-EARTH GEOPHYSICS

The UMP was proposed in 1960. It began in 1962 and ended in 1970 and included a wide range of geophysical, geochemical, and geological studies related to problems of the earth's interior. Special attention was devoted to continental margins and island arcs; the world rift system; and the viscosity and mechanical behavior of the mantle. At the beginning of the UMP, these subjects had received little attention. By the end of the UMP, it was evident that they were, indeed, the critical scientific areas related to fundamental problems of the earth's interior.

One of the more ambitious aspirations of the Upper Mantle Committee was to determine whether continental drift had really occurred. The subject of intense debate in the 1920's, continental drift had fallen into disrepute for apparent lack of a mechanism. Paleomagnetic evidence in the 1950's revived the subject, but the theory won few converts, for there was still no convincing mechanism. In 1960, the year in which the UMP was proposed, Harry Hess put forward the concept subsequently called sea-floor spreading. It suggested a mechanism and rationale for the lateral motion and opened the way to a theory of global tectonics. However, at the time of its introduction, evidence to support the concept was extremely scant.

The wide range of research connected with the UMP led to a remarkable accumulation of evidence that supported the concept of sea-floor spreading. For example, new bathymetric data showed that the continental margins of the Atlantic constitute a better fit even than the well-known fit of the shorelines; geochronological studies showed the fit of isochrons across the South Atlantic; paleomagnetic reconstruction could be made of ancient continental positions; regularity and symmetry of magnetic lineations were found around most of the midocean ridges; the earth's surface was shown to be divided into a small number of plates that move relative to one another; and high heat flow was measured at the midocean ridges (consonant with uprising convection currents).

By the end of the UMP, sea-floor spreading had been transformed from an imaginative insight by Hess to a hypothesis, then to a theory, and, in the minds of most solid-earth scientists, to an established fact. This development constitutes a revolution in earth sciences comparable to the revolution that took place in physics in the early part of this century.

The ability of the new concepts of global tectonics to explain an enormous range of geological and geophysical information, including practical questions about mineral and petroleum deposits, gave impetus to a new international program, the Geodynamics Project. This is an international program of research on the dynamics and dynamic history of the earth, coordinated by the Inter-Union Commission on Geodynamics (ICG), which was established by ICSU at the request of IUGG and IUGS.

The Geodynamics Project deals principally with movements at the surface and within the upper portions of the earth's interior. If one considers that the outer shell consists of a number of lithospheric blocks, then the program can logically be divided into four major parts: movement of these blocks relative to each other; movements beneath the blocks that have an effect on them; movements, primarily vertical, within the blocks; and past movements of blocks (not necessarily of the present configuration) as reflected in the geological record. Each of these parts will serve as a focus for special working groups.

7.4 OCEANOGRAPHY

The IGY was the beginning of worldwide international cooperation in oceanographic research at sea, although exchange of ideas and results from research had existed for many years. Much was accomplished in both physical and chemical oceanography and in studies of the sea floor and the suboceanic crust. Definitive velocity measurements were made of two major currents in the tropical Pacific. The Cromwell Current, or Equatorial Undercurrent, was found to be a thin ribbon of water flowing at high and nearly constant speeds due east beneath the surface for thousands of miles along the Equator. The Equatorial Countercurrent, north of the Equator, was shown on one occasion to extend from the surface down to depths of the order of 1000 m and to fluctuate rather widely at any one location in both speed and direction. In the Atlantic, the existence of a countercurrent under the Gulf Stream was demonstrated by direct velocity measurements from U.S. and British ships.

Wave-recording instruments and tide gauges were set up in remote locations in order to study worldwide seasonal changes in sea level and long-period waves on the surface of the sea.

The shape of the East Pacific Rise, the Pacific counterpart of the Mid-Atlantic Ridge, was defined by detailed echo-sounding profiles across its axis. The data showed that this great spine, which extends from Antarctica to Lower California, is intersected by transverse ridges, reaching out from South America and extending into the central Pacific, which may be older than the Rise itself. Near the axis of the Rise, measurements of the heat flow from the interior of the earth gave values four to eight times higher than the average for the oceans as a whole, while extremely low values were obtained near the Peru–Chile Trench and in a broad area west of the Rise. This pattern of heat flow is one of the underlying lines of evidence for the hypothesis that the mantle of the earth is slowly turning over in giant convection cells that drag the plates of the earth's crust with them.

Continuous measurements of atmospheric and oceanic carbon dioxide on a worldwide basis were begun during the IGY in an attempt to determine the earth's carbon dioxide budget and especially to find what is happening to the carbon dioxide produced by the burning of fossil fuels. This program has continued, and it is now clear that carbon dioxide in the atmosphere is increasing by about 0.7 part per million per year–0.25 percent of the present atmospheric content. One third to one half the carbon dioxide produced by fossil fuel combustion becomes dissolved in the oceans.

A Special Committee of the International Council of Scientific Unions, the Scientific Committee on Oceanic Research (SCOR), was formed in 1958 to continue the cooperation begun during the IGY. The SCOR's most important early accomplishment in fostering international cooperative research was the initiation and planning of the International Indian Ocean Expedition, in which some 20 countries and 40 ships participated. Measurements made during this expedition established the shape, structure, and compartmentation of the ocean basin, delineating the topography of the complex midocean ridge system and establishing its continuity with similar seismically active rises in the central Atlantic and the southeastern Pacific. One branch was shown to run aground in the Gulf of Aden–Red Sea–African Rift region. Magnetic profiles in the northwest Indian Ocean provided early and convincing support for the idea of spreading of the sea floor and creation of new crust at midoceanic ridge crests.

In contrast to the awesome trenches of the Pacific or the classical midocean ridge of the Atlantic, the Indian Ocean seems best characterized by the extensive aseismic submarine plateaus or rises of intermediate depths and intermediate crustal structure, neither truly continental nor truly oceanic. Such "microcontinents" may be litter abandoned, or flotsam carried along, in the breakup and drift of large continental masses, or perhaps the tracks or scars along which displacement has occurred. One such block, containing the granitic Seychelles Islands, is of truly continental character and is 600 million years old while nearby, perhaps in fault contact, is sea floor less than a sixth as old.

Measurements of the equatorial currents showed striking differences from those of the Pacific, largely because of the monsoonal wind system. Although an equatorial undercurrent is present, it is dependent on the phase of the monsoon, being found only at the end of the winter monsoon. During the summer monsoon, subsurface currents along the equator are weak and variable. At the same time, along the western boundary of the

Arabian Sea, the Somali Current flows northward at speeds exceeding 6 knots. The strong poleward winds lead to intense upwelling on the Somali and Arabian coasts, bringing food to the surface layers and providing the basis for a rich fishery resource. Only a few months later, with onset of the winter monsoon, flow along the Somali coast is reversed and replaced by a weak drift toward the south.

An extension of cooperation in oceanography is being developed by the Intergovernmental Oceanographic Commission, together with SCOR and other scientific groups. An international network of measuring buoys and bottom-mounted devices throughout the world ocean is being planned to monitor changes in ocean conditions and to help in predicting such changes. The organization, coordination, and operation of such a system, together with its associated satellites, ships, and aircraft, will require a greater and more effective effort on the part of international scientific institutions, both governmental and nongovernmental, than ever before.

7.5 METEOROLOGY

During the IGY, ICSU and WMO worked to increase the coverage of meteorological soundings, including extensions in altitude to 100,000 ft; to establish three meridional chains to study transport from pole to pole; and to conduct studies of such atmospheric constituents as CO_2.

The joint effort of many nations has made possible meteorological work that would otherwise have been impossible. For instance, for many years, the concept of a direct and purely gravity-driven atmospheric circulation over the great ice domes of Antarctica and Greenland—that is, the idea of the "glacial anticyclone"—had dominated meteorological thinking. However, IGY observations at stations and during traverses on the high plateau of the southern continent yielded a different picture. The surface winds are directly related not only to the configuration of the terrain but also to the strength of the temperature inversion in the lowest few hundred meters of the atmosphere. Over the perennial ice fields, this inversion persists during most of the year, with a vertical increase of temperature by 15 to 30°C from February through November. Above the surface inversion, cyclonic circulation of upward increasing intensity prevails in the entire troposphere all year long.

Polar studies, IGY rocket soundings, the pioneering infrared and cloud-cover satellites of the IGY, and the general pattern of cooperation provided methods and tools that bore fruit after the IGY. They led, for example, to the TIROS, TOS, and Nimbus series of experimental and research satellites and to the operational satellites that followed. They now make possible, in conjunction with the numerical modeling and computer innovations of von Neumann and his colleagues, the formulation of an attack on global atmospheric processes, GARP, which is discussed also in Chapter 5.

7.6 POLAR RESEARCH

Problems of logistics in the Arctic regions had been largely solved before the IGY. The Soviets in particular, because of obvious geographical interests, had conducted sea and air operations along their coast and far into the Arctic Basin to study sea ice and water as well as the lower and upper atmospheres, and they had established observatories on ice floes. In the United States, the Air Force Cambridge Research Laboratories had pioneered in ice-flow work, notably in the discovery and use of Fletcher's Ice Island, which served as a scientific station for many years. Studies in the Arctic were intensified during the IGY and have continued to the present, albeit at a lower level on the part of the United Sattes.

The Antarctic, however, was another matter. It was far more remote, inaccessible, and uncongenial; nor was there national interest linked to sovereign lands, coasts, and coastal waters. Yet the planners of the IGY considered both polar regions of appreciable scientific interest, for these were vast zones replete with unique regional problems, and, in any case, a global view of the earth required their investigation. Distinctive features of unique scientific interest are present: The magnetic poles control the configuration of the aurora and the geometry of processes in the magnetosphere; and these cold regions influence the dynamics of flow in the oceans and the atmosphere. Because some 70 percent of the earth's fresh water is in the form of Antarctic ice, better measurement of this ice is pertinent to an understanding of the earth's heat and water budget.

Largely because of logistical difficulties encountered by early explorers and investigators, scientific knowledge of Antarctica and the surrounding waters was scanty at the advent of the IGY. During the IGY, however, these difficulties were overcome by the use of ski-equipped air transports, helicopters, snow tractors, and icebreakers, which enabled 12 nations to maintain 54 winter stations and many more summer stations. Data gathered at these stations and the hundreds of ship stations in Antarctic waters revolutionized the knowledge of Antarctica in several disciplines.

Representatives of six nations worked together on synoptic weather analysis and research programs at the IGY Antarctic Weather Center at Little America. They provided regular analyses for routine operational purposes and supported other scientific disciplines in which meteorological information was needed. They also developed programs of research into the basic synoptic processes of Antarctica and the Southern Hemisphere. Another element of international cooperation was the vigorous program in which scientists of several nations spent periods of time ranging from a summer to a year or more working at the scientific bases of other nations. These exchanges covered meteorology, biology, geology, upper-atmosphere studies, and the like and are still being conducted.

Two of the general consequences of the IGY Antarctic program were the political agreement on Antarctica and the creation of an ICSU mechanism for post-IGY cooperation in Antaractic research. The first of these took the form of the Antarctic Treaty, ratified in 1961, which put aside national claims, dedicated Antarctica to peaceful scientific uses, and provided for inspection of any station by representatives of any signatory nation. A Treaty Organization was established, meeting biennially, which relies on the ICSU committee mentioned below for scientific counsel. It is generally acknowledged that the Treaty—the first in history to devote a region, admittedly barren and forbidding, to peaceful purposes—was possible only because the IGY had established a suitable amity among national participants.

The second consequence of such cooperative investigations was the establishment by ICSU of the Special Committee for Antarctic Research* (SCAR), which has continued to foster cooperative studies in the Antarctic regions. The number of fixed stations has decreased from the peak of the IGY, but the number of disciplines and the degree of mobility have increased. Although some solid-earth and biological studies were included in the IGY, these were not its primary subjects. The SCAR has encouraged substantial efforts in marine biology (as well as physical oceanography) in Antarctic waters and has devoted markedly increased attention to studies of flora and fauna and of geology, while continuing the investigations of interface and atmospheric sciences.

7.7 SPACE PHYSICS

Interests in the conduct of *in situ* measurements had led the organizers of the IGY to advocate the use of instrumented rockets and satellites. Significant results accrued—for example, the discovery of the Van Allen radiation belt, precursor to the present knowledge of the magnetosphere. Interest in this area led to the formation of the Committee for the International Year of the Quiet Sun (IQSY).

The IQSY was an international cooperative program of observations in

* Now the *Scientific* Committee for Antarctic Research.

solar-terrestrial physics. The IQSY was designed to complement, during the solar minimum of 1964–1965, the IGY programs in sun-controlled disciplines that had been carried out during the solar maximum of 1957–1958. A Special Committee* of ICSU was established in 1962, with representatives of the four unions with interests in this field (IAU, IUPAP, IUGG, and URSI), two representatives from COSPAR, with which it worked very closely, and delegates from the 71 participating nations. Agreements were made about the establishment or continuation of nets of observing stations, the collection and distribution of synoptic data, the geographical and time distribution of work, and coordination with existing programs (such as space programs). The Special Committee for the IQSY continued in existence until the end of 1967, to bring the results of the 1964–1965 observing program to an orderly conclusion by, for example, publishing the seven-volume *Annals of the IQSY*.[12] Some national committees (including the U.S. National Committee on IQSY) also published reports.

During the IGY, an international network of strategically located cosmic ray neutron monitors was set up that, on a large scale, made use of the earth's magnetic field as a magnetic spectrometer to study energy dependence, directional dependence, and time dependence of energetic solar and galactic particles. These techniques yielded the first understanding of particle propagation in interplanetary space and thereby of the physical processes taking place in the interplanetary medium. It became clear that diffusion processes govern this propagation; the existence of the solar wind was postulated, and the structure of interplanetary magnetic fields was deduced. These ideas were confirmed by the first earth satellites and deep-space probes. The time dependence of the cosmic ray intensity does not depend on the presence of the earth; the responsible mechanism is heliocentric in character. Direct measurements of the stream of solar plasma confirmed the existence of a solar wind, and the interplanetary magnetic fields were found to be consistent with those predicted. Within the few years covered by IGY and IQSY, the physics of interplanetary space developed from an almost unknown field to a well-established one.

Illustrative of work fostered by the IQSY are satellite studies of the sun. During the interval covered by the IQSY, several SOLRAD's recorded solar x-ray emissions in several wavelength bands. They transmitted a continuous telemetry signal, which could be received and decoded by anyone with adequate equipment. As a result, atmospheric scientists and aeronomers in many countries built inexpensive receivers and obtained data in real time

* A "Special Committee" of ICSU is one on which national scientific groups, as well as unions, are officially represented.

on solar x-ray emission, which they could then use to correlate with their own ground-based measurements. Solar radiation data from other spacecraft (such as the NASA Orbiting Solar Observatory, or OSO, series) and terrestrial observatories were also (and still are) distributed so that correlations of this same kind can be performed after the fact.

The understanding of the emission and propagation of energetic particles, typically produced in solar flares, was fragmentary in the pre-IGY years but gained enormously through synoptic observations from many disciplines. Solar patrol and observations by a worldwide net of solar observatories in the optical region and in various bands of radio emission continuously covered the solar disk. These observations were combined with the detection of solar x rays and with measurements of particle fluxes and spectra by a variety of means. Satellite detectors, balloon-borne detectors, neutron monitors, and polar-cap radio absorption and forward-scattering measurements provided data on measurable aspects of the phenomenon. As a result, there exists today a detailed picture of particle propagation between sun and earth and of the relative roles played by anisotropic diffusion, convection, and acceleration in particle transport.

The Proton Flare Project, proposed by IAU Commission 10 and endorsed by the IQSY Committee, was organized at the end of the IQSY. At that time, the level of solar activity was beginning to increase, improving the chances of occurrence of a proton flare. Many types of observation were involved: time histories and radio-frequency spectra of the entire range of electromagnetic radiation, from radio frequencies to x rays; particle emissions in all accessible energy ranges, from both spacecraft and ground observatories; and time series of maps of the active region's magnetic fields. The effects of the proton flare on the ionosphere and the geomagnetic field also were included. A system for issuing quick warnings of imminent or occurring solar flares was established, and July 1966 was selected as one of the alert periods. As a consequence, a large body of observational material was obtained on the event of July 7, 1966, and on other lesser flares occurring during the following several days. It was possible to deduce (with some disagreement over details) the mechanism by which the complex magnetic fields in the flare region broke down, with the release of energy in the form of high-energy particles (up to several GeV) and intense x rays, and to trace the effects of these emissions in the interplanetary medium and their interaction with the earth's magnetosphere and atmosphere right down to the ground. This was the first occasion on which such a complete history of the interlocking phenomena related to a single solar proton flare had been assembled and interpreted. (The collected results are published in Volume 3 of the *IQSY Annals.*)

Both the IGY and IQSY showed the value of the international interdisciplinary approach, and, as a consequence, considerable momentum built up in solar-terrestrial physics. With the close of the IQSY, it seemed desirable to establish new machinery at the international and interunion level to provide for continued planning, coordination, and information and data exchange. The Inter-Union Commission on Solar-Terrestrial Physics was established in 1966 for these purposes, with representatives from the same international organizations as the IQSY Committee, although its contacts with national bodies are more informal.

7.8 COMMITTEE ON SPACE RESEARCH

Before the close of the IGY, it was clear that a mechanism for international cooperation in space research would be useful; thus it was that ICSU organized the Committee on Space Research. Space research involves many different disciplines in the life and physical sciences; therefore, the Council of COSPAR consists of representatives of 11 unions as well as 35 nations. Working groups in relevant areas afford opportunities for discussion of important space research problems; symposia and publications provide for the discussion and dissemination of results; some data interchange has come about through the World Data Center System (see below); and cooperative programs have been realized—for example, COSPAR coordinated the space aspects of the IQSY and later programs in solar-terrestrial physics. NASA first offered bilateral international collaboration through COSPAR, and its subsequent program has involved several scores of nations in cooperative rocket investigations, the flight of experiments on U.S. space systems, and a variety of training and educational relationships.

A necessary, if small, role of COSPAR is that of maintaining an international roster of objects launched into space, just as the IAU keeps track of comets and asteroids. This information, together with the particulars of space experiments suited to international cooperation (for example, satellites with radio beacons that can be used for ionospheric studies and real-time telemetry from satellites monitoring solar x rays), allowed nations that did not have satellite-launching capabilities to participate in space research.

Another COSPAR project that has contributed significantly to both physical science and constructive international collaboration is the International Reference Atmosphere, a set of tables modeling the high atmosphere on the basis of *in situ* information from rockets and satellites.

The annual COSPAR meetings and symposia have provided an opportunity for space scientists from all nations to intermingle, to become acquainted personally as well as through the scientific literature. In such a fast-moving field, the pace of publication through normal channels is too slow, and personal communications influence future experiments long before the earlier results appear in print. COSPAR has also provided an invaluable scientific link between east and west, short-circuiting the additional delays of translation of journal articles.

7.9 EXCHANGE OF DATA

Early in the planning of the IGY, its organizers recognized that, however successful its planning and execution might be, the program would remain incomplete without an exchange of original data and of calibrated, summary tables. The exchange of summary tabular data was realized in the publication of the *Annals of the IGY* (48 volumes). The exchange of original data was achieved through the World Data Center system of three international centers: WDC-A in the United States, WDC-B in the Soviet Union, and WDC-C, distributed by discipline among several nations in Europe, Australia, and Japan.

The contents of the centers vary with the discipline. Under WMO auspices, complete meteorological data of the standard parameters is augmented by special studies begun during the IGY. Samples taken at agreed intervals are stored for other fields, such as ionospheric physics. Complete interchange of all magnetograms takes place. These centers, by international agreement, are maintained under the auspices of ICSU to handle data gathered since the IGY. Details of the types of data to be exchanged are considered periodically by groups of specialists, such as COSPAR, the Upper Mantle Committee, or the Inter-Union Commission on Solar-Terrestrial Physics. Groups that interpret the needs of the scientific community and the growing areas of practical applications oversee the data-exchange programs.

Data frequently requested, such as solar-wind measurements and geomagnetic indices, are published. Detailed data, such as hourly average intensities of the neutron component of cosmic rays, are stored on magnetic tape. Other data, such as photographs of comet tails, are held at the originating observatory, but the Data Centers serve as guides to what data exist and where they are. The data-exchange mechanism takes a heavy burden of correspondence off institutions and individual scientists and assures ready and rapid access to needed data.

8 Characteristics of Earth and Planetary Physics and Priorities in This Subfield

A discussion of the role of physicists, their funding, and their productivity in the earth and space sciences presents a number of aspects that are not encountered in the traditional core fields of physics. Consequently, the data appearing in this chapter must be viewed in the proper context.

There are two major sources of difficulty. First, earth and planetary physics is not a separate, clear-cut field. Physics is only one of the disciplines involved in the earth and space sciences. Manpower requirements and funding, priority, and planning decisions relate to major science activities; rarely, if ever, do they pertain specifically to the role of the physicist. To be effective, the physicist working in earth and space sciences must combine his efforts with those of engineers, chemists, life scientists, meteorologists, geologists, and others. In doing so, his view of himself and the research process can change significantly from that engendered by his training and experience as a physicist. Possibly, he will cease to identify himself as a physicist, and, as the statistics will show, he probably will not publish in recognized physics journals or in any journal covered by *Physics Abstracts.*

There are others who, in the broadest sense of the word, are physicists but probably have never considered themselves as such. For example, a typical pattern that could well be found in a department of physics is an undergraduate major in applied mathematics, followed by graduate studies that include many advanced courses in physics and mathematics and thesis research in dynamical meteorology. However, a student following this program would not be considered part of the physics community.

The second difficulty relates to the inhomogeneity of the earth sciences. Although all have a similarity in approach, the differences are such that mobility among the different areas of specialization generally is not great and statements that cover the entire field are likely to be heavily weighted by the unique characteristics of one particular component. Thus the geological sciences employ by far the largest number of scientists and differ from other earth and space sciences in having a big industrial segment and a high proportion of teachers. The geological sciences constitute a mature field, transfers to and from which are slow. Meteorology also is a relatively mature field with low turnover of personnel, but industrial involvement is small and that in government-supported activities, high. In meteorology a significant component of research and development tends to be conducted by non-PhD's. On the other hand, both aeronomy and interplanetary research have always been populated largely by physicists. Data on earth and space scientists in the societies belonging to the American Institute of Physics (AIP) tend to be weighted toward the characteristics of this group. Oceanography has only recently emerged as an important field for physicists. It has a large interface with the life sciences. Indirect costs, in the form of ship time, for oceanographic research tend to be disproportionately large compared to those in the geosciences and atmospheric sciences. Both oceanographic and space sciences employ small numbers of scientists compared to the atmospheric sciences, which, in turn, are dwarfed in comparison with the geological sciences.

An additional difficulty is that research supported by the National Aeronautics and Space Administration (NASA) has unique characteristics, and, since these expenditures are very large, they distort the fiscal statistics for research in the earth and space sciences. Since NASA activity is recent in inception and is subject to rapid change, it is probably responsible for much of the mobility that now occurs in earth and space sciences. Aeronomy and interplanetary physics have been strongly supported by NASA, and these fields display characteristics peculiar to the NASA mission. Finally, NASA has large in-house research activities and also has supported institute-type research on university campuses. Many of those working in such institutes can be identified as physicists; therefore, statistics on the nature of employment at universities in the earth and space sciences tend to show a relatively large proportion of nonteaching appointments.

8.1 MANPOWER

8.1.1 *The Interface with Physics*

This section presents manpower data culled from a variety of sources that employ classifications sometimes differing from those adopted in this report. This Panel's major concern is the relationship between the discipline of physics on the one hand and the earth sciences and related space sciences on the other. The interface defines the earth and planetary physics subfield. The relationship to physics usually is established by (a) the nature of the highest academic degree, (b) AIP membership, and (c) self-identification. All three methods of identification are fallible; in particular, all can fail to identify a man whose research methods are those of physics but whose graduate department was not physics.

Other disciplines also can be identified by means of society membership: geosciences by the American Geological Institute (AGI), and atmospheric sciences by the American Meteorological Society (AMS). Oceanography and space sciences have no universally accepted national scientific societies;

all the earth and space sciences are represented in the American Geophysical Union (AGU). In each of these societies, some members identify themselves as physicists. It is probable that such people also are represented in the AIP and respond to its questionnaires. Thus, one finds physicists classified in the earth and space sciences and earth and space scientists classified as physicists.

Degree patterns from studies[13] of the 1967 annual meetings of the AGU and the AMS provide data on the overlap between physics and the earth and space sciences. In a sample of 183 authors of presentations made at the 48th Annual Meeting of the AGU in April 1967, the field of highest academic degree indicated by 33 percent was physics; 1 percent indicated astronomy. Of a random sample of 416 from all of the attendants at this meeting, 27 percent also indicated physics as the field of highest academic degree, with 2 percent reporting their highest degrees in astronomy. The principal work activity indicated by those in both samples was basic research; four fifths of the authors and about two thirds of the attendant group ranked this activity first or second in terms of demand on their time.

A similar study conducted at the 47th Annual Meeting of the AMS in January 1967 and at the meteorological sessions of the April 1967 Annual Meeting of the AGU showed that 17 percent of the sample of 130 authors had received their highest academic degrees in physics and 1.5 percent in astronomy. Physics was the field of highest academic degree of 11 percent of the 281 meeting attendants sampled, and astronomy that of 7 percent. Basic and applied research were the most time-demanding work activities in both these groups, with the meeting-attendant sample also reporting substantial allocations of time to administrative work.

These studies collected information only on highest academic degree and provided no data on those with terminal degrees in geophysics or meteorology who held undergraduate or master's degrees in physics.

Another view of the interface between earth and planetary physics and physics is afforded by the Section membership of the National Academy of Sciences (NAS). Sections are chosen by members to correspond to their primary field of interest. The Geophysics Section (39 members) currently includes all the areas subsumed under earth and planetary physics, although a few active members in the subfield have chosen the Sections of Physics (109 members), Geology (41 members), Astronomy (36 members), and Applied Sciences (28 members). These data are difficult to interpret, but they suggest that a significant proportion of those NAS members who regard themselves as being within the broad discipline of physics are primarily concerned with the subfield of earth and planetary physics.

The most recent edition (1968) of the National Science Foundation

(NSF) publication *American Science Manpower* provides data on those portions of the National Register of Scientific and Technical Personnel concerned with members of the AMS and the AGI. These data, together with 1968 and 1970 data on physicists identified with earth and planetary physics in the National Register survey, afford a partial description of the characteristics of scientists working in the subfield.

American Science Manpower 1968 (NSF 69-38) shows that approximately 10 percent of the 298,000 participants in the 1968 National Register Survey were identified with atmospheric, geological, and oceanographic activities and that about 11 percent were identified with physics. (If it is assumed that some 85 percent of the PhD and non-PhD populations taken together respond to the Register Survey, then in 1968 the total U.S. scientific manpower–Register respondents plus nonrespondents–in the atmospheric, geological, and oceanographic disciplines was about 34,000, and the total in physics, about 38,000.) About 1600 of the physicists work in atmospheric, geological, or oceanographic disciplines; however, very few of those in the earth and space science groups work in physics. Table IX.5 presents data from *American Science Manpower* and the National Register Survey for 1968 on earth and space sciences manpower. The table gives some indication of the involvement of physicists in earth and space sciences (see the column showing the number identified with the AIP). A higher percentage of PhD's, composed largely of physicists, is found in the space sciences than in atmospheric, geological, and oceanographic disciplines taken together. In the space sciences, the percentage of PhD's (49 percent) is, as might be anticipated, close to that for physics (44 percent). Only 16 percent of the 23,160 scientists identified with atmospheric, earth, and ocean sciences in the 1968 Register Survey held PhD's.

Table IX.5 also illustrates the imbalance among the different fields that comprise the earth and space sciences. Geological sciences dominate numerically. Atmospheric sciences also account for a large part of the scientific manpower. Space and oceanographic sciences are relatively smaller. However, these smaller areas have a large fraction of PhD's.

8.1.2 *Employment Patterns*

Table IX.6 compares the employment patterns of earth and planetary physicists with patterns for all physicists. Earth and planetary physicists were less involved in teaching than physicists in general and more involved in government research. As subsequent discussion will show, the teaching involvement of university-based earth and planetary physicists is low, and the contrast with physics is stronger than the figures in Table IX.6 indicate.

TABLE IX.5 Manpower in Earth and Space Sciences[a]

Field of Employment	No. of Scientists	Principal Processing Society[b]	No. Identified with AIP	No. Holding PhD	% of Total Represented by AIP Scientists	% of Total Holding PhD
Atmospheric sciences	5,232	AMS	453	639	8.7	12.2
Structure & dynamics	1,127		378	435	33.6	38.6
Other	4,105		75	204	1.8	5.0
Geological sciences	17,198	AGI	197	3,040	1.1	17.7
Geophysics	2,873		171	431	6.0	15.0
Other	14,325		26	2,609		18.2
Oceanography	730	AGI	53	203	7.3	27.8
Total atmospheric, geological, and oceanographic sciences	23,160		703	3,882	2.7	15.7
Space science	816	AIP	626	399	76.7	49.0

[a] Data are based on *American Science Manpower 1968* (NSF 69-38) (Table A-48) and on the 1968 National Register of Scientific and Technical Personnel.

[b] Society responsible for the collection and processing of data in a particular discipline as a part of the National Register effort. If a respondent's form is returned to a society, yet he seems to belong more appropriately to another society, his form is sent to that society for review and incorporation in their returns if it meets their qualifications. This sorting among the societies involved in the Register effort since 1962 has refined to some extent professional identification of respondents.

TABLE IX.6 Employment Patterns of Earth and Planetary Physics PhD's and Non-PhD's Compared with Those for the Total Physics Population[a]

Employing Institution	Earth and Planetary PhD's, $N = 712$ (%)	All Physics PhD's, $N = 16{,}631$ (%)	Earth and Planetary Non-PhD's, $N = 673$ (%)	All Physics Non-PhD's, $N = 19{,}705$ (%)
College and university	38	51	33	28
Industry	23	23	20	33
Government	22	9	31	15
Research center	13	12	4	6
Other	4	5	12	18

[a] Data are from the 1970 NSF National Register of Scientific and Technical Personnel.

8.1.3 *Work Activities*

Table IX.7 summarizes the primary work activities of earth scientists and earth and planetary physicists. The differences in work patterns are striking. A high proportion of PhD's and non-PhD's in earth and planetary physics are engaged in research (67 percent and 64 percent, respectively). The comparable figures for research for all physics are 55 percent for PhD's and 47 percent for non-PhD's. In the earth sciences as a whole, the figures are 32 percent (PhD) and 18 percent (non-PhD). This finding reinforces the conclusion, derived from almost all types of data on these fields, that patterns of activity and employment in earth and planetary physics differ from those for both the earth sciences as a whole and physics.

This is true also of the roles and responsibilities of PhD physicists and earth scientists in universities. In the earth sciences as a whole, some four fifths (82 percent) of those employed in universities hold faculty appointments. Only half (51 percent) of the earth and planetary physicists have faculty appointments. The percentage for physics falls between these extremes.

8.1.4 *Mobility*

A comparison of the data from the 1968 and 1970 National Register surveys shows changes in the earth and planetary subfield during this two-year interval. The PhD population of the subfield increased by 25 percent, from 567 in 1968 to 712 in 1970. (The number of PhD's in earth and planetary physics in 1964 was 309.) The nondoctorate population increased from 599 in 1968 to 673 in 1970, a 12 percent growth. About 71 percent of the net growth in the subfield resulted from new PhD's. The median age of the

TABLE IX.7 Primary Work Activities of Earth Scientists[a]

Degree and Field	Work Activities (%) Research[b]	Teaching	R & D Management	Other[c]
PhD's				
Earth and planetary (N = 707)	67	13	15	5
AGI scientists (N = 4689)	30	42	9	19
AMS scientists (N = 493)	50	23	16	10
Non-PhD's				
Earth and planetary (N = 657)	64	7	11	18
AGI scientists (N = 21,746)	17	18	6	59
AMS scientists (N = 5413)	20	5	7	67

[a] Data for the American Geological Institute and the American Meteorological Society respondents are based on *American Science Manpower 1968;* those on earth and planetary physics are based on the 1970 National Register survey.

[b] Research includes basic and applied as well as the design and development category, which has a negligible effect in these groups.

[c] The "other" category includes exploration and forecasting, which are major activities of many scientists in the geological and meteorological groups.

earth and planetary physicists changed from 37.8 years in 1968 to 37.4 years in 1970. The median age of new PhD's was (surprisingly) 31.5 years.

In addition to an influx of new PhD's, there was a large turnover of manpower in earth and planetary physics between 1968 and 1970. Only 55.6 percent of those working in this subfield in 1970 reported working in it in 1968. One third (33.6 percent) of the 1968 subfield population apparently left. The age distribution of those moving from the subfield does not differ greatly from that of those remaining in it.

Other aspects of mobility are movements between employers and geographical movements. In 1970, in contrast to 1968, 20 percent of earth and planetary physics PhD's were new PhD's, 13 percent had changed employer, 28 percent were new to the subject but had kept the same employer, while 39 percent had continued in the subject with the same employer.

Geographic mobility provides only a lower bound for the various kinds of movement. Earth and planetary physicists remained in the same location more frequently than did other physicists. Although 33 percent of the 1968 PhD's shifted from the subfield, only 6.7 percent moved to another geographic location.

Personal experience of members of the Panel on Earth and Planetary Physics suggests that the findings on mobility may not be typical of current conditions. A major group in earth and planetary physics has been supported by NASA. This group accounts for a disproportionate number of research workers in universities who do not have faculty appointments. Retrenchment in federal programs probably will influence this group of physicists, and future statistics will no doubt reflect the impact of government policy.

Another change in mobility patterns could result from the change in attitude of young physicists. There is an increasing desire to apply the skills of physics to the solution of national and social problems. Earth and planetary physics exercises an attraction in this context, evidence for which is an upsurge in applications for postdoctoral positions in this subfield. Obviously, this trend is amplified by decreasing research opportunities in the traditional subfields of physics.

Earth and planetary physics has not been able to respond to this surge of interest. As subsequent sections show, the subfield has been characterized by, at best, level budgeting in spite of increasing numbers of employees, the rapid growth of knowledge, and the exciting prospects for its exploitation. Moreover, the subfield is relatively mature; manpower levels and employment patterns cannot be changed rapidly to meet new opportunities.

These statements can be illustrated by an example—applications for postdoctoral fellowships at the National Center for Atmospheric Research (NCAR). Table IX.8 shows the total number of applications and awards, with the number of physicists who had never before been connected with atmospheric research ("physics applicants" and "physics awards") shown separately. The response that the Center has been able to make to the pressure from outside is small. Of the six physicists awarded fellowships in 1971-1972, four have been sent to the University of Chicago to gain experience in atmospheric problems.

Possibly, this situation is transitory. The emotional climate toward environmental issues could change, and universities may modify their policies

TABLE IX.8 Postdoctoral Applications and Awards at the National Center for Atmospheric Research

Applicants and Awards	1968–1969	1969–1970	1970–1971	1971–1972
Total applicants	26	46	129	123
Physics applicants	9	17	43	40
Total awards	10	13	11	18
Physics awards	1	2	0	6

toward graduate education in physics. However, there is at present a rare opportunity, perhaps of short duration, to obtain a transfusion of new skills into earth and planetary physics.

8.2 PRODUCTIVITY

Perhaps the least satisfactory comparison with other subfields is provided by the publication data obtained from a study of 13 issues of *Physics Abstracts* (numbers 856, 860, 861, and 869–878). Members of the Panel on Earth and Planetary Physics often cannot find their own publications in the *Abstracts* because the editors do not classify them as physics. Moreover, the subfield is characterized by a regrettably large "gray" literature of observatory and project reports, which are known and available to those who work in the field. Although the Panel does not necessarily endorse such a manner of disseminating research findings, its significance as a means of publication cannot be overlooked.

The total sample of articles from U.S. institutions in all subfields of physics taken together was 1183, of which 637 were experimental in nature and 546 theoretical. About 10 percent of the papers dealt with the subject matter of earth and planetary physics. Of these 117 papers, slightly more than half (53 percent) were experimental. Table IX.9 shows the types of institutions in which the reported work was conducted. Most of the research, both experimental and theoretical, took place in universities, with important contributions from government laboratories and industry.

An additional survey of physics theses abstracted in *Dissertation Abstracts* during the first six months of 1970 showed that 5.5 percent (33) of the 597 theses listed in the physics section were concerned with the subject matter of earth and planetary physics. Theses that appeared in the engineering section of *Dissertation Abstracts* also were examined, and 870 were identified that dealt with material that might legitimately have ap-

TABLE IX.9 Institutions in Which Published Research in Earth and Planetary Physics Was Conducted

Nature of Work	Institutions			
	University	Government	Industry	Research Center
Experimental	35	15	7	5
Theoretical	26	12	11	6
TOTAL	61	27	18	11

peared in the physics section. Of these 870 physics-oriented engineering theses, 55 (6 percent) treated subjects within the scope of earth and planetary physics.

According to this last statistic, more activity in earth and planetary physics may take place under the aegis of engineering than under that of physics. This conclusion is not entirely surprising. In Chapter 4, we drew parallels between the relationship of physics to the earth sciences in education and the relationship of physics to other fields such as engineering and engineering physics. In many cases, the simplest solution to problems created by the rigid departmental structure of universities has been to develop branches of the earth sciences in, or in combination with, engineering schools.

This factor emphasizes the difficulty of collecting adequate data on physics and the earth sciences and interpreting such data as are available. The data used throughout this section in manpower pertain to the recognizable core area of physics.

8.3 FUNDING

Under the heading Environmental Sciences, federal documents identify expenditures as Atmospheric, Oceanographic, and Earth Sciences; however, there is no subdivision into earth and planetary physics. As a result, the Panel had to rely on general experience with the subfield to arrive at rough estimates of expenditures.

The discussion of manpower suggests that a large proportion of those engaged in research in environmental sciences are identified with earth and planetary physics. In addition, an unusually high proportion of the earth and planetary physicists, as compared to other physicists and to other scientists working in the environmental sciences, were engaged in research. It seems evident that there are few areas of basic or applied research in the environmental sciences in which education in physics is not profitable. Consequently, in one sense, the entire research expenditure in the environmental sciences is potentially related to earth and planetary physics and we shall not distinguish sharply between them in this discussion. (The following section of this chapter on "Activities in Individual Disciplines" offers a more detailed discussion.)

Table IX.10 presents an overview of federal expenditures on research and development by department for fiscal year 1970. The figures are further broken down into physics and environmental sciences, showing that these two fields account for more than 50 percent of the basic research expenditure in NASA, the NSF, the AEC, and the Departments

TABLE IX.10 Some Characteristics of Research and Development Obligations of Federal Agencies, Fiscal Year 1970 (estimated)[a]

					Percent of Basic and Applied Research Budgets Allocated to Physical and Environmental Sciences			
					Basic		Applied	
Agency or Department	Total R&D Obligation ($millions)	Basic Research (%)	Applied Research (%)	Development (%)	Physical Sciences	Environmental Sciences	Physical Sciences	Environmental Sciences
Agriculture	280.1	41	56	3	18	–	13	–
Commerce	86.0	34	45	22	40	35	11	49
Defense	7756.0	3	15	82	27	20	17	8
Interior	226.0	26	50	23	16	45	9	14
Transportation	359.0	5	18	77	–	60	–	–
AEC	1347.6	21	10	69	80	–	47	–
NASA	3776.7	20	18	62	45	33	–	14
NSF	294.2	87	8	5	31	21	–	15

[a] Data are based on *Federal Funds for Research, Development, and Other Scientific Activities. Fiscal Years 1969, 1970, and 1971*, Volume XIX (NSF 70-38).

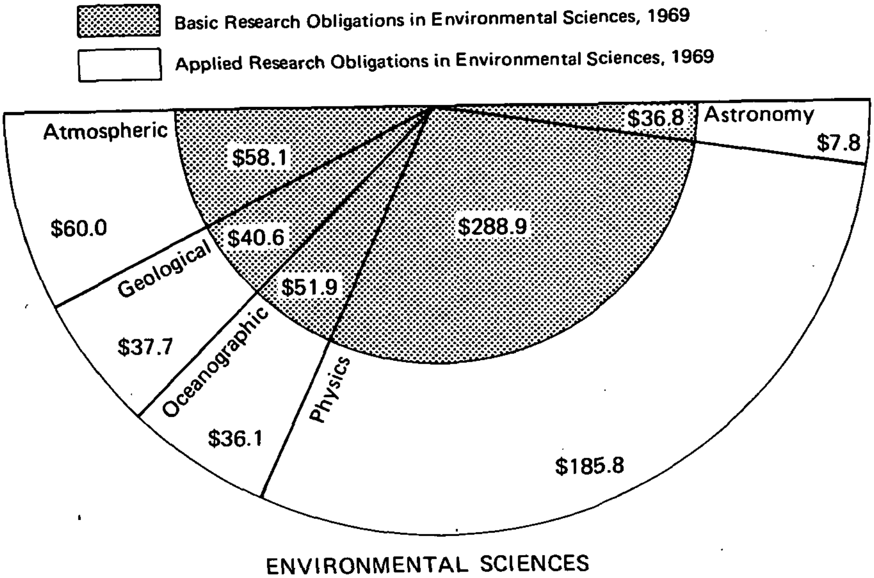


FIGURE IX.14 Federal obligations (in millions of dollars) for basic and applied research in environmental sciences in 1969 (not including NASA obligations). [Source: *Federal Funds for Research, Development, and Other Scientific Activities. Fiscal Years 1969, 1970, and 1971,* Volume XIX (NSF 70-38).]

of Transportation and Commerce. The breakdown also indicates that total expenditures in the environmental sciences do not lag far behind those for physics. A comparison of research expenditures in physics, astronomy, and the environmental sciences appears in Figure IX.14. The data in this figure do not include NASA expenditures.*

Table IX.11 shows the breakdown of funds for basic research in the atmospheric, geological, and oceanographic components of the environ-

* NASA expenditures on Space Science and Applications (OSSA budget) are treated in a subsequent section. Unless these data are examined in detail, they distort the total support picture; for example, NASA research obligations in the environmental sciences were \$474 million in 1966, as compared to \$291 million in all other agencies, and \$258 million in 1969, as compared to \$293 million in other agencies.

To judge by the breakdown in Table IX.10, NASA budgets should have a similar effect on physics expenditures. However, it is not clear that this is the case. The NASA expenditure on "space physics" amounted to about \$50 million in 1969, and most of this expenditure is properly considered under earth and planetary physics. Thus it is difficult to understand how 45 percent of NASA's expenditure could be attributed to physics. The fiscal situation becomes clearer if NASA expenditures are separated from those of other agencies.

TABLE IX.11 Support of Basic Research in the Environmental Sciences in 1969[a]

Agency or Department	Percent of Total Basic Research Support	Support ($millions)			
		Atmospheric	Geological	Oceanographic	Total
NSF	32.3	24.0	9.5	15.1	48.7
Interior	16.1	–	23.1	1.2	24.3
Army	4.0	4.5	1.0	0.5	6.0
Navy	20.4	5.6	1.7	23.5	30.8
Air Force	11.2	15.4	1.5	–	16.9
Commerce	6.5	7.2	1.5	1.1	9.8

[a] Data are based on *Federal Funds for Research, Development, and Other Scientific Activities. Fiscal Years 1969, 1970, and 1971,* Volume XIX (NSF 70-38).

mental sciences in some of the federal agencies (excluding NASA) that allocate substantial support to such work. Not unexpectedly, the Navy is heavily engaged in the support of basic research in oceanography, and the Department of the Interior (that is, the U.S. Geological Survey) supports much basic research in geology. It is surprising to note how little the Environmental Science Services Administration (ESSA), now absorbed into the National Oceanic and Atmospheric Administration, of the Department of Commerce, invests in research. Basic research in the atmospheric sciences is supported principally by the NSF and the Air Force.

The most significant fact about the federal support of basic research in the environmental sciences is that it has been relatively level during recent years, despite a steady increase in the number of scientists involved. Table IX.12 presents these data, beginning with 1966. (Prior to that date, NASA expenditures were not separately identified in the NSF tabulations.) The 1971 estimates included in the table predated congressional hearings on appropriations for that fiscal year and are typically high. The data show that, in spite of pressures from Congress and the public for more effort in relation to environmental problems, there has been no increase of expenditure in the past five years. Since inflation has been continuous during this interval, the actual level of effort must have declined. A more detailed examination of work in various fields of the environmental sciences in the following section of this chapter reflects this same level pattern.

One might expect that industry would support a substantial amount of research in geophysics, but this does not seem to be the case. The NSF publication *Research and Development in Industry, 1967* (NSF 69-28)

TABLE IX.12 Federal Basic Research Expenditures in the Environmental Sciences ($millions)[a]

	Basic Research Expenditures by Fiscal Year					
Science	1966	1967	1968	1969	1970 (est.)	1971 (est.)
Atmospheric	57.3	59.8	58.1	58.1	59.7	66.7
Geological	32.6	37.6	40.2	40.7	43.1	43.7
Oceanographic	60.3	39.2	51.7	51.6	53.6	60.7
TOTAL	150.2	136.6	150.0	150.4	156.4	171.1

[a] Data in the table do not include NASA expenditures. Data are based on *Federal Funds for Research, Development, and Other Scientific Activities. Fiscal Years 1969, 1970, and 1971,* Volume XIX (NSF 70-38).

indicates a total of $20 million for support of basic research in "geology, geophysics, and other earth sciences" in 1967 in a table that includes both federal and industrial funds. The amount supplied by industry from its own resources appears to be small. Although no specific data are available for oceanographic and atmospheric sciences, industrial expenditures in these sciences also are probably small.

8.4 ACTIVITIES IN INDIVIDUAL DISCIPLINES

8.4.1 *Atmospheric Sciences*

During the 1960's, the atmospheric sciences in the United States underwent major changes both in research expenditures, which rapidly increased, and in outlook. Prior to 1960, meteorological research was poorly supported, with emphasis on short-term projects and with relatively little input from the physics and mathematics communities. Aeronomy, which had separated almost completely from meteorology in the early 1950's, had relatively close connections with physics, particularly in countries other than the United States.

The interests of a small group, including J. von Neuman, E. Teller, L. Berkner, and C. Rossby, led to a new appraisal of the status of meteorology. In the National Academy of Sciences–National Research Council report, *Research and Education in Meteorology: An Interim Report,*[14] the Committee on Meteorology emphasized that the subject offered far greater opportunities than had been exploited, particularly in regard to the use of physical research. It was proposed that the basic research budget be doubled from $5 million to $10 million and that a National Center for Atmospheric Research be created. The former suggestion led to an Atmospheric Sciences Program in the NSF, with the National Center for Atmospheric Research at Boulder, Colorado, becoming the largest single expenditure of this program.

Further projections appeared in *The Atmospheric Sciences 1961–1971.*[15] This report recommended a total federal research budget of $605 million and a production rate of 180 PhD's per year in 1971. Data presented in Figure 13 indicate a total 1969 basic and applied research expenditure in atmospheric sciences of about $118 million; there has probably been no significant increase since 1969. The PhD production is difficult to evaluate; however, some 80 new PhD's in the atmospheric sciences now apply for NCAR postdoctoral fellowships. According to some estimates, the total production of PhD's in the atmospheric sciences is about 120 per year.

Since the creation of the NSF program and the founding of NCAR, there have been three major thrusts in research: the World Weather Program, weather and climate modification, and environmental quality.

The Global Atmospheric Research Program (GARP) and the World Weather Watch (WWW) together constitute the World Weather Program, to which the United States has a strong national commitment. In 1961, President Kennedy proposed to the United Nations "further cooperative efforts between all nations in weather prediction and eventually in weather control." In May 1968, the 90th Congress approved a resolution stating that "it is the sense of Congress that the United States should participate in and give full support to the world weather program."

International activity, including U.S. participation, in meteorological operations has always been at a high level (see Chapter 7). The reasons for currently increased activity are the new opportunities that have emerged as a result of the availability of satellites and large computers.

The objectives of GARP, as stated by the World Meteorological Organization–International Council of Scientific Unions Joint Organizing Committee are as follows:

The Global Atmospheric Research Programme is a programme for studying those physical processes in the troposphere and stratosphere that are essential for an understanding of:

(a) The transient behaviour of the atmosphere as manifested in the large-scale fluctuations which control changes of the weather; this would lead to increasing the accuracy of forecasting over periods from one day to several weeks.

(b) The factors that determine the statistical properties of the general circulation of the atmosphere which would lead to better understanding of the physical basis of climate.

This programme consists of two distinct parts, which are, however, closely interrelated:

(i) The design and testing by computational methods of a series of theoretical models of relevant aspects of the atmosphere's behavior to permit increasingly precise description of the significant physical processes and their interactions
(ii) Observational and experimental studies of the atmosphere to provide the data required for the design of such theoretical models and the testing of their validity.

According to the plan for U.S. participation in GARP,[16]

Three requirements have been identified which are essential to meeting the objectives of GARP:

(1) The development of a global observing capability
(2) The availability of electronic computers with speeds at least 100 times faster than the speeds of the most powerful computers in use today
(3) The conduct of regional field programmes and computer modelling experiments to improve the physical and mathematical basis of long-range prediction.

The GARP is a research program of limited duration. To the extent that it is successful, its results will be used to update the entire international meteorological effort. The second phase is the WWW, a WMO program. The following reports offer discussions of the feasibility of GARP and WWW from the scientific, technical, and manpower points of view:

Atmospheric Exploration by Remote Probes (NAS-NRC, 1969)
The Feasibility of a Global Observation and Analysis Experiment (NAS-NRC Publ. 1290, 1966)
Educational Implications of the Global Atmospheric Research Program (NAS-NRC, 1969)
The Global Atmospheric Research Programme (GARP) (Report of the Study Conference held in Stockholm, 28 June to 11 July, 1967. International Council of Scientific Unions–International Union of Geodesy and Geophysics–Committee on Space Research–World Meteorological Organization)

These reports show that major technical and scientific problems must be solved if a 10- to 14-day forecast capability is to be achieved. The solution will be difficult, even in an extensive and strongly motivated research effort. Active debate over some aspects of these problems could continue until the end of this century.

These reports do not include cost estimates, but it is probable that the directly identifiable cost of GARP, allocated over a number of years, will be about $200 million. NASA has a growing interest in this program. Because of its long-term character and close relationship to human needs, it could have an important stabilizing influence on NASA science budgets.

To assess accurately the extent of NASA's involvement in meteorology is difficult. Planned programs include TIROS and improved TIROS satellites together with Nimbus E–G and follow-up efforts. The results from these satellites are needed to plan for GARP, which is a separate commitment. A Synchronous Meteorology Satellite also is planned. All of these satellites form part of the Earth Environmental Sciences program of NASA, which is scheduled to increase from $133 million in 1971 to a high level of $393 million in 1977. From 25 percent to 50 percent of this expenditure could be for meteorological programs. Thus the directly identified expenditures on GARP are likely to be supplemented with large-scale support by NASA if GARP and WWW are to achieve their goals.

The costs of WWW are undetermined at the present time. If GARP achieves its objectives, it could lead to a relatively firm assessment of returns in relation to expenditure in meteorology. Caution in planning probably is advisable. With computer and satellite hardware available, it is relatively easy to initiate programs of massive cost. Very little work has been done on the problem of cost effectiveness. The most widely quoted work on this problem is *The Value of Weather*, by Maunder.[17] This book is an ambitious effort by a single author, but the information available to him is fragmentary. Benefit–cost ratios from 10 to 100 are claimed with existing systems, and the author concludes "that increased expenditure for weather forecasting services is an investment that will reap rich dividends." This conclusion cannot be seriously questioned. Nevertheless, in the light of the human tendency to exaggerate benefits and minimize costs, at least when data allow different interpretations, and considering the step-function increase in expenditure that could result for WWW, it would be advisable to subject this view to closer scrutiny.

Modification and control of climate and weather are two of the ultimate goals of meteorological research. Progress toward these goals is discussed in *Weather and Climate Modification: Problems and Prospects.*[7] This report describes the many different attempts to modify weather phenomena and reports some positive results. It also discusses accidental modification; this problem currently is receiving widespread attention as part of the atmospheric pollution problem. The need for a single agency to exercise overall control in this complex area is suggested.

Perhaps the major conclusion of the report is that one cannot with impunity tamper with such a complex system as the lower atmosphere without a complete understanding of the fundamental physical and chemical processes involved. This depth of knowledge is still lacking. Therefore, the report strongly urges increased effort in fundamental investigations. Larger computers are needed, and, for this purpose alone, an increase of funds from $5 million (1965) to $30 million (1970) was recommended. Rapid advance in control and modification of weather may not take place until the objectives of GARP have been achieved.

In the 1970's, an increased effort in meteorology will be required to deal effectively with air pollution problems. To the extent that sources of pollution are known, the most direct approach is political and legal, with the goal of modifying effluent in quantity or quality. Meteorology can play a part in the solution of problems of ventilation, smokestack design, and the like. The solution of more difficult problems through meteorological research is apt to be a long-term effort requiring increased knowledge of many aspects of atmospheric behavior. As is true of problems of modification and control of weather, the completion of the GARP effort is prerequisite to rapid advance.

The response of funding agencies to the great changes in scientific activity in the atmospheric sciences was discussed earlier in this chapter; another representation appears in Figure IX.15. The upper two sections of this figure depict NASA activity, which will be discussed in a separate section. The cost of aeronomy is small in these programs. Most planetary work to date has been concerned with atmospheres and is related to meteorology and aeronomy, but again it is best to separate NASA expenditures from those of other agencies.

If meteorology and meteorological satellites are considered together, a rapid rise in support between 1960 and 1963 is apparent, followed by almost level funding (that is to say, decreasing capability) from 1964 to 1970. Table IX.13 presents the same picture in regard to the NSF budget (the chief source of support for basic research outside NASA). In meteorology, only GARP shows increasing support, rising perhaps to $1.5 million in 1970. Aeronomy (that is, the program as a whole) declines. The increase in facilities is needed to take care of budget items that were dropped by Defense agencies, for example, the Arecibo and Jicamarca observatories; there is a net loss on a national basis.

The optimism in regard to NCAR is indicated by the budget projections to 1976 that appear in Figure IX.16. This trend contrasts strongly with the support to universities from the federal government. Recent figures on university support issued by the Interdepartmental Committee for Atmospheric Science make possible a comparison of averages over the three years 1965, 1966, and 1967 with those for 1969, 1970, and

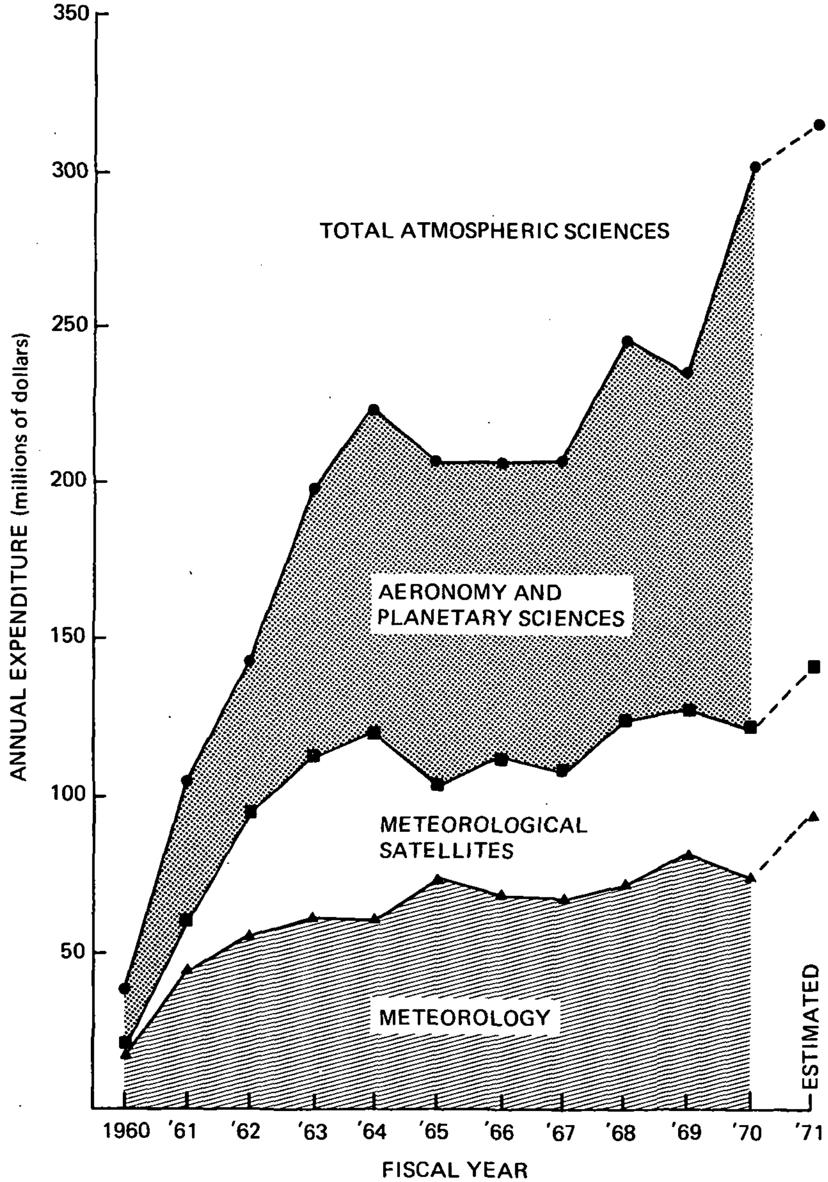


FIGURE IX.15 Federal expenditures for atmospheric sciences basic research during the 1960's. Total expenditures during 1960–1969: aeronomy and planetary atmospheres, $875 million; meteorology, $509 million; meteorological sciences, $406 million. [Source: *The Atmospheric Sciences and Man's Needs: Priorities for the Future.*[7]]

TABLE IX.13 The NSF Atmospheric Sciences Program, 1966-1971 ($Millions)[a]

Program	1966	1967	1968	1969	1970	1971[b]
Meteorology.						
Weather modification	2.0	3.0	3.0	2.5	3.0	3.0
Meteorology	3.4	3.8	4.0	4.1	4.0	4.5
GARP	0	0	0.2	0.5	1.5	1.9
TOTAL	5.4	6.8	7.2	7.1	8.5	9.4
Aeronomy						
Aeronomy	1.5	1.9	1.9	1.5	1.7	2.1
Solar-Terrestrial	1.6	1.4	1.8	1.9	2.3	2.4
IQSY	1.7	0.7	0	0	0	0
TOTAL	4.8	4.0	3.7	3.4	4.0	4.5
Facilities and Miscellaneous	1.1	1.0	0.8	0.3	0.2	2.5
NCAR[c]	9.2	12.6	11.4	10.2	10.2	13.6
TOTAL	20.5	24.4	23.1	21.0	22.9	30.0

[a] A fraction of the expenditure of the Office of Polar Programs (approximately $10 million per annum) is used for atmospheric sciences. Breakdowns of this program are not available, but the expenditure on atmospheric work seems to be level.

[b] 1971 figures are estimates.

[c] Other NSF figures show NCAR costs up to $1 million higher per year. NCAR also receives some contract support from NASA, which is not shown in the table.

1971 (not yet realized and perhaps too high). During these few years, total expenditures fell from $40.5 million to $35.2 million, while inflation additionally decreased the effectiveness of the investment by 20 percent to 40 percent.* The NSF investment in the universities increased during this interval, but the AEC, DOD, and NASA decreased their budgets. The resulting picture is one of withdrawal of support for atmospheric science by large agencies, while the NSF does not obtain the increase needed to assume the entire university support. (See also Appendix IX.C.)

A curious paradox in the NSF funding in atmospheric sciences exists,

* Clearly, inflation is a crucial factor in judging the support of this and other programs in earth and space sciences. The Panel is indebted to Harvey Brooks for the following opinion: "The general GNP deflator considerably underestimates the inflation problem, particularly in a field like earth and planetary physics, with a recent injection of very young people. The existence of a nonsteady-state age distribution creates an inflation of professional salaries which is at least twice the rate of inflation of salaries in general. I would be very much surprised if the rate of inflation in aeronomy and planetary physics is less than eight percent per year, [it was] probably more like 10 percent to 12 percent during the last two years."

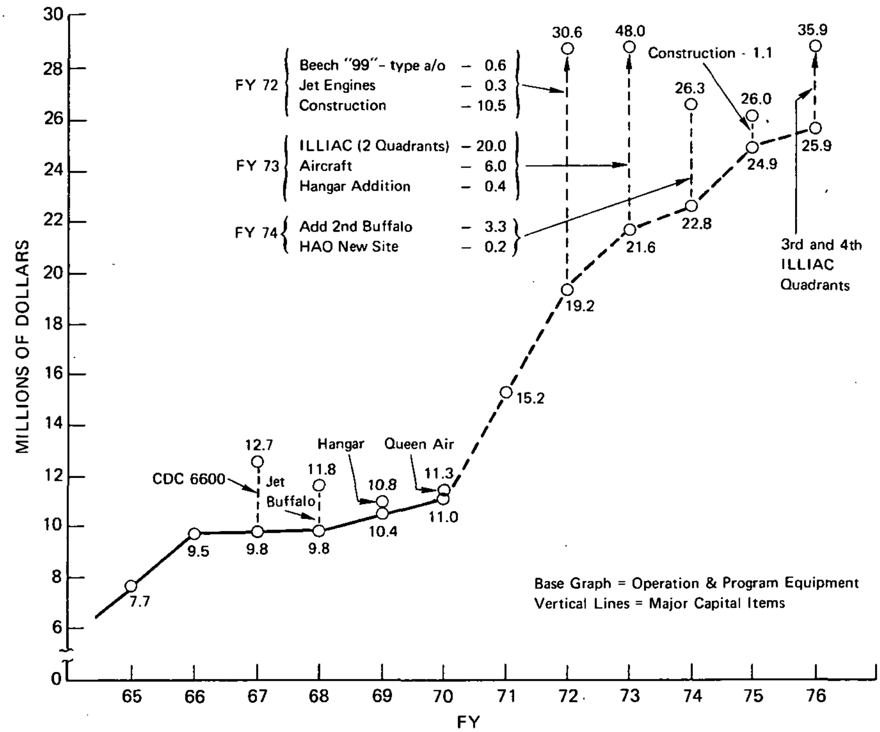


FIGURE IX.16 NCAR appropriations, fiscal years 1965–1970, with projections to fiscal year 1976.

one that is repeated in other earth sciences. A subject, with steadily increasing opportunities and research personnel, has received decreasing support in the universities in terms of fixed dollars during the years 1965–1970. Under the circumstances, one might anticipate a steady increase in proposal pressure on the NSF, but the data in Table IX.14 suggest the opposite. The dollar requests by individuals have *decreased* with time, while the amount awarded has *increased*. We must conclude that during this five-year period either university investigators deliberately restrained their requests for support or that the figures reflect administrative and bookkeeping changes within the foundation. Foundation officials have no simple explanation for the paradox.

The first footnote in Table IX.14 indicates radical change in funding patterns; during the first six months of fiscal year 1971, requests for support increased 100 percent in dollar amount over 1970. There were no funds available to meet this influx of good proposals. According to

TABLE IX.14 Individual Proposals Submitted and Supported under the NSF Program in the Atmospheric Sciences for Fiscal Year 1965 through Fiscal Year 1970[a]

Proposals Received and Granted	1965	1966	1967	1968	1969	1970
Total new proposals received ($millions/yr)	12.5	19.1	19.1	17.0	17.9	16.9
Total proposals granted[b] ($millions/yr)	4.2	6.5	7.2	7.7	7.9	7.9
Percent of total dollars granted	34	34	37	45	44	47

[a] During the first six months of fiscal year 1971, the number of proposals received was 60 percent greater than for the same period in fiscal year 1970, and 100 percent more money was requested.
[b] Figures do not include weather modification, facilities, or NCAR.

the NSF program directors, there is no difference in quality between the new applications and those previously submitted. The new applications came from people of the highest competence who had recently lost support from DOD, NASA, and other agencies. The trends in support considered in this section have not yet reflected this development, which suggests that past data will be misleading as a basis for future projections.

Research priorities in the atmospheric sciences were proposed by a recent (1971) report of the NAS-NRC Committee on the Atmospheric Sciences entitled *The Atmospheric Sciences and Man's Needs: Priorities for the Future.* The objective of the study was to identify feasible research programs that could contribute most to important human needs within the next five to ten years and to propose an ordering of research priorities to meet these needs. The report enumerated four major objectives, each supported by recommendations of programs to achieve the objectives. The recommended priorities represent a moderate but significant shift in research emphasis and in patterns of federal support in the atmospheric sciences.

The four major objectives proposed for the next decade were

1. To extend the capability for useful prediction of weather and atmospheric processes
2. To contribute to the capability to manage and control the concentrations of air pollutants
3. To establish mechanisms for the rational examination of deliberate and inadvertent means for modifying weather and climate
4. To reduce substantially human casualties, economic losses, and social dislocations caused by weather.

Recommendations under Objective 1 considered phenomena of all time and space scales. Global prediction was given highest priority, followed by a proposed pilot project to determine the practical usefulness of modern technology in providing local weather information. Research on mechanisms of climate change came next, followed by research on mesoscale weather (periods of 2 to 12 hours). Recommendations under Objective 2 supported model development and measurement programs for urban pollution, establishment of a simple global monitoring system, and study of the chemistry of precipitation and dry fallout. Recommendations under Objective 3 addressed particularly the administrative and public policy problems that are emerging as critical in this field. A need was identified for continuing attention at the level of the Executive Office to the public policy aspects of weather modification, the National Oceanic and Atmospheric Administration was proposed as best equipped for the principal administrative responsibility in this field, and a resolution to the United Nations General Assembly was urged that would dedicate all deliberate weather and climate modification efforts to peaceful purposes. Other recommendations were addressed to research activities. Objective 4 was the focal point of the report and was supported by all prior recommendations.

Estimates of research expenditures required to implement these recommendations appear in Table IX.15. The figures are totals for the decade and are incremental to existing expenditures. A total increase of \$453 million, or \$45.3 million per year on the average, is proposed. This represents an average annual increase for the decade of about 7 percent in meteorology and meteorological satellites.

The Atmospheric Sciences and Man's Needs: Priorities for the Future did not attempt to review priorities in aeronomy or planetary atmospheres; however, aeronomy has been treated in two reports of the Geophysics Research Board: *Physics of the Earth in Space: The Role of Ground-Based Research*[18] and *Upper Atmosphere Observatory, Criteria and Capability.*[19] The second of these reports is an amplification of one proposal made in the first for a new incoherent backscatter facility at $L = 4$ on the Canadian border for the investigation of ionospheric and magnetospheric dynamics. The cost is estimated at between \$12 million and \$15 million.

The report on the role of ground-based research is concerned with continuing support for a field of high scientific achievement and interest and relatively low cost. Although no branch of earth and planetary physics is divorced from or even distant from application, the report does not discuss human needs and the influence of these considerations on priorities. However, it states that the low cost of research and the inter-

TABLE IX.15 Implications of the NAS Committee on Atmospheric Sciences (CAS) Recommendations for New Research Funding in the Atmospheric Sciences during 1970–1979

Atmospheric Sciences Program	NAS-CAS Recommendations	NAS-CAS Priority Ranking	New Money ($millions) For Research Operations	For Research Facilities
Weather prediction	Full implementation of GARP	1-1	75	Computer facilities*
	Local weather watch	1-2	4	8
	Dynamics and modelling of climate	1-3; 3-3	25	Computer facilities*
	Mesoscale research	1-4	15	10
Air quality	Urban models and experiments	2-1	30	Computer facilities*
	Remote sensing and research	2-1; 3-5	25	13
	Global air quality	2-2	20	Computer facilities*
	Regional models and measurements	2-3	16	Computer facilities*
Weather and climate modification	Assignment of direction and administration of program	3-1; 3-2; and 3-6	–	–
	Weather modification laboratory	3-4; 1-4	35	8
	Cloud dynamics and physics models and measurements	3-5; 1-4	25	Computer facilities* 4
*Consolidated computer requirements	In support of 1-1, 1-3, 2-1, 2-2, 2-3, 3-3, and 3-5	–	–	140
TOTALS	–	–	270	183

esting phenomena involved have made this field an important point of contact between conventional physics and the earth and space sciences.

The recommendations include new permanent facilities, greater use of existing facilities or approaches, modernization of facilities or techniques, new types of instrumentation, and new mobile stations. Unfortunately, it is not likely that new support will become available. NASA, DOD, and NOAA are steadily decreasing their support in this field. The NSF budgets generally have held firm, but the competitive demands on them increase, particularly in the form of requests to sustain large facilities from which other agencies have withdrawn support. The report draws particular attention to the value of joint use of ground-based and space techniques. NASA support for this branch of space physics will be considered in a later section of this chapter.

8.4.2 *Geological Sciences*

The geological sciences (geology and geophysics) are by far the largest component of the earth and space sciences. Figure IX.17 gives present and projected data on manpower in these sciences. The figure does not include a category of "other earth scientists" consisting of oceanographers, hydrologists, geochemists, and other unidentified groups. The total number of geologists and geophysicists in 1970 was nearly 28,000; with "other earth scientists," this total would rise to 37,000.

Figure IX.18 shows distribution of geologists and geophysicists by academic degree and employment, based on data collected by the AGI. The high percentages employed in the petroleum industry are apparent.

The aspirations of the geophysical sciences community are discussed in the NAS–NRC report *Solid-Earth Geophysics: Survey and Outlook.*[20] A more recent survey appears in the report of an *ad hoc* committee of

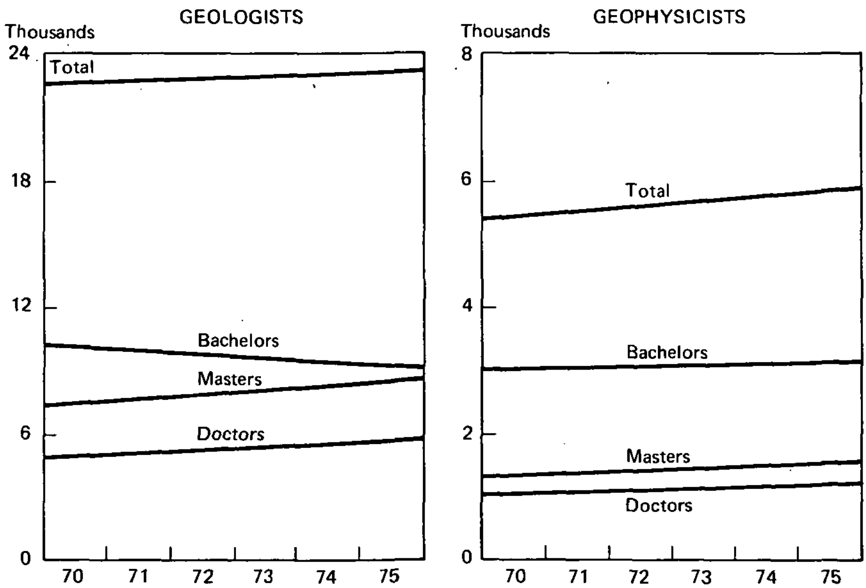


FIGURE IX.17 Projected employment in geology and geophysics. [Source: Committee on Manpower, *Manpower Supply and Demand in Earth Science* (American Geological Institute, Washington, D.C., 1971).]

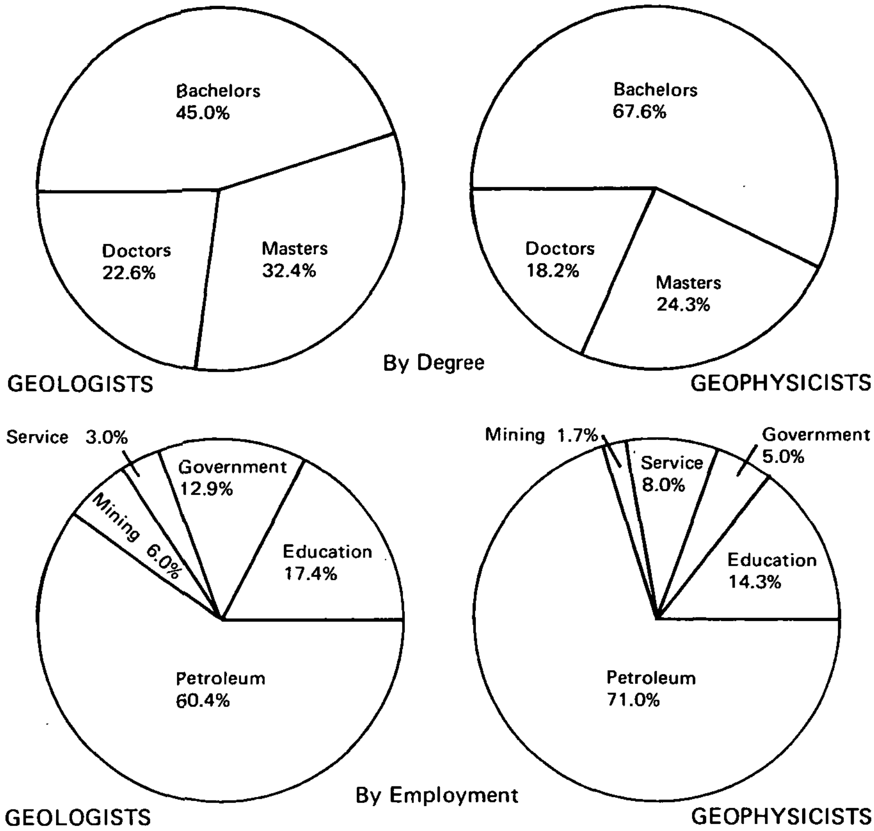


FIGURE IX.18 Distribution of geologists and geophysicists by degree and employment. (Based on data collected by the American Geological Institute, 1970.)

the International Union of Geodesy and Geophysics–International Union of Geological Sciences on *Long-Range Program of Solid Earth Studies.*[21] These reports refer infrequently to applied problems, although applications are important considerations even for the most fundamental research.

On the other hand, *Earthquake Engineering Research*[22] deals with short-term and highly practical questions, and *Seismology: Responsibilities and Requirements of a Growing Science*[23] is concerned with the entire range of research in one branch of geophysics.

The impact of NASA in this field is likely to be large, as this agency seeks long-term programs, such as GARP, with direct applications to man's needs. A report describing the possibilities for NASA and assigning priorities is *The Terrestrial Environment: Solid-Earth and Ocean Physics.*[24] The emphasis of this document is on geodesy.

The early review, *Solid-Earth Geophysics: Survey and Outlook,* was written at about the time of inception of the International Upper Mantle Project and the AEC VELA-Uniform Program. Its recommendations for a ten-year science program at a level of about $50 million per year do not take account of the recent rapid advances in the knowledge of earthquakes and motions of the mantle.

Both, *Earthquake Engineering Research* and *Seismology: Responsibilities and Requirements of a Growing Science* pose specific questions. The former suggests an investment of $38 million per year over a ten-year period for engineering purposes alone and points out that the construction industry investment rate in highly seismic areas of the United States is $10 billion per year. The proposed research investment is trivial in financial terms, even without consideration of the lives that could be saved.

The report on seismology estimates a somewhat similar level of expenditure for all aspects of the subject: $503 million over a ten-year period. This report is concerned with aspects of seismology other than construction and the protection of human life, that is to say, oil and mineral prospecting and nuclear test detection. It identifies the need to convert the World-Wide Network of Standardized Seismographic Stations into a national facility of greater general availability.

As in meteorology, the effect of NASA spending on solid-earth geophysics will be great if the projected expansion of applications programs comes about. Under earth physics, the proposed NASA program lists multidisciplinary earth observatories, precision position satellite, magnetic survey, GEOS, ATS, and a number of DRAGSAT and SEASAT satellites. Details are not defined, but it appears that 25 percent of the earth environmental sciences budget ($133 million in 1971, rising to $393 million in 1977) could be connected with solid-earth geophysics in some degree.

In general, federal support has not responded to the existing proposals by the science and engineering communities, and these have not yet been updated to reflect the excitement that now pervades the geophysical community as a result of the extraordinary advance in the knowledge of the basic processes shaping this planet: sea-floor spreading, plate tectonics, and global tectonics. A quotation from *Geodynamics Project: Development of a U.S. Program*[25] conveys some of this feeling:

Five years ago no one would have predicted that micropaleontologists, geomagnetists, marine geomorphologists and seismologists would be working together to supply the crucial test of a concept comparable to that of the Bohr atom in its simplicity, elegance, and ability to explain a wide range of diverse observations. . . . [T] he result has been a revitalization of the earth sciences comparable to that which swept physics at the beginning of the century.

The research expenditures shown in Table IX.12 give little indication of these momentous developments. Geological sciences receive somewhat less federal research support than atmospheric or oceanic sciences and are essentially level or decreasing in capability. Federal support is illustrated by the development of the NSF geological sciences program. Prior to 1958, there was virtually no NSF funding. The level of support rose to $2.28 million in 1961. An increase to $5.217 million in 1963 was associated with the inception of the Upper Mantle Project. From 1963 to 1966, the program grew at a modest rate, but from 1966 to 1970, the budget [excluding the Joint Oceanographic Institutions Deep Exploratory Survey (JOIDES)] did not change (see Table IX.16) and the extent of the scientific effort in individual programs has been sharply decreased by inflation.

Table IX.17 compares geophysics proposals received with proposals granted. The trends in the support of geophysics are almost identical to the pattern in the atmospheric sciences (see Table IX.14).

The footnote to Table IX.16 shows that a substantial part of the activity in geophysics has been supported by other sections of the NSF budget. This is a possible explanation of the lack of unusual pressure in the form of new proposals in the main geophysics program. The 60 per-

TABLE IX.16 The NSF Geological Sciences Program 1966-1971 ($millions)[a]

Fiscal Year	Program Allocations					
	Geology	Geochemistry	Geophysics	Total	JOIDES	Total
1966	1.86	2.94	2.73	7.53	5.40	12.93
1967	1.27	3.18	3.49	7.94	0.0	7.94
1968	1.46	2.97	3.28	7.81	4.10	11.91
1969	1.34	3.36	3.22	7.92	2.50	10.42
1970	1.42	3.07	3.36	7.85	6.60	14.45
1971	–	–	–	7.80	7.2	15.00

[a] In addition to these figures, $1.08 million and $3.25 million for geology and geophysics, respectively, are assigned under the oceanography program as parts of the International Decade of Ocean Exploration expenditure for 1971. Geological projects are also supported under the Office of Polar Programs. In addition to IDOE expenditures, a substantial part of the NSF Oceanography Program is devoted to geological projects ($3.18 million in 1969 and 1970) (see Table IX.18).

TABLE IX.17 Proposals Received and Supported by the NSF Geophysics Program for Fiscal Year 1965 through Fiscal Year 1970[a]

Proposals Received and Granted	1965	1966	1967	1968	1969	1970
Total new proposals received ($millions)	4.9	7.4	7.3	8.6	7.3	6.9
Total proposals granted ($millions)	2.6	2.8	3.5	3.3	3.3	3.5
Percent of total dollars granted	54	38	48	38	46	50

[a] Estimates for FY 1971 indicate a 60 percent increase in proposal load, which is almost an exact parallel to the atmospheric sciences program.

cent increase in proposal pressure noted in the first six months of fiscal 1971 also occurred in the atmospheric sciences program. (In fact, the phenomenon occurred in almost every NSF program.) However, the geophysics program differs from the atmospheric sciences in that the 1971 budget shows no increase over 1970 (see Table IX.13). Therefore, the geophysics program must redistribute existing resources to meet new demands.

On the whole, the geological sciences have fared slightly worse than atmospheric sciences and oceanography. The reason might be traced in part to certain historical characteristics of the field. Formerly, and traditionally, it required only a small "entrance fee" to do research in geophysics. The field is populated by scientists who are used to total independence and to undertaking relatively small projects. The atmospheric and oceanic sciences, on the other hand, have long been accustomed to the idea that large cooperative efforts are needed, and they have developed a mix of small projects and large programs, each of which supports the other.

Geophysics suddenly has been presented with a new situation as a result of the recent developments in plate and global tectonics. However, the solution is not as simple as devising a single large project. Typically, the cost of key research projects connected with recent developments tends to be between $0.2 million and $2.0 million—too small for a large program but too large for a small project.

The objective of the Geodynamics Project, sponsored by ICSU, is to devise programs that will resolve this dilemma and permit progress commensurate with the extraordinary scientific opportunities. The concept of the U.S. National Committee is to synthesize a program from existing elements by identifying major omissions, for example, a high-pressure facility for large samples. Many intermediate-sized programs thus can come together to form what constitutes a national enterprise without concentrating the effort in a few institutions. The climate of cooperation between members of the profession is advancing to the point at which such a communal enterprise could be effective.

8.4.3 *Ocean Sciences*

Several recent reports indicate that a substantial effort in oceanography is timely. Four reports on this subject are

President's Science Advisory Committee, *Effective Use of the Sea* (U.S. Government Printing Office, Washington, D.C., 1966).

Committee on Oceanography, *Oceanography, 1966: Achievements and Opportunities*, Publ. 1492 (National Academy of Sciences–National Research Council, Washington, D.C., 1967).

Committee on Oceanography and Committee on Ocean Engineering, *An Oceanic Quest: The International Decade of Ocean Exploration*, Publ. 1709 (National Academy of Sciences–National Academy of Engineering–National Research Council, Washington, D.C., 1969).

Commission on Marine Science, Engineering, and Resources, *Our Nation and the Sea: A Plan for National Action* (U.S. Government Printing Office, Washington, D.C., 1969).

These reports suggest that oceanography in the United States lags in almost all respects, particularly in requirements for new ships and facilities.

An Oceanic Quest discusses these shortages in the context of the International Decade of Ocean Exploration (IDOE), a program to which the United States was committed by the President on March 8, 1969. The decade is intended to extend throughout the 1970's.

There is a wide range in estimates of costs for this international program—from \$1 billion to \$5 billion over the decade. Much of this expenditure is for systematic studies of the ocean bottom, including deep-drilling operations; the physical, chemical, and dynamical properties of the ocean at levels from the surface to the abyss; the interactions between air and ocean; and living organisms in the sea. These surveys, like most investigations in the earth sciences, are of value as strictly empirical compilations of data. However, they also will shed light on existing physical theories and direct attention to new phenomena for which physical explanations are required. As much as half the U.S. contributions to IDOE could be concerned with physical oceanography.

Our Nation and the Sea is a broad analysis of a total national program that resulted from two years of work by the Commission on Marine Science, Engineering, and Resources. Although its proposals are based on marine requirements, they also advocate the linking of meteorology and oceanography in a single two-fluid system. The central proposal was for the creation of a new agency, the National Oceanic and Atmospheric Administration (NOAA), reporting directly to the President. The report recommended that this agency be created by combining four existing agencies, five existing programs (three from the NSF, including NCAR), and six new programs. Its objective was to strengthen the national capability in a wide variety of different ways, using industry, universities, and private and state research programs as well as in-house activities. The proposals include the programs of the IDOE but are much more extensive.

At the time *Our Nation and the Sea* was issued, it was estimated that the agencies and programs to be included in NOAA would represent an expenditure of roughly $800 million annually and that this expenditure should gradually increase to $2 billion per year by 1980. This amount would not represent the total federal expenditure for oceanic and atmospheric work since other agencies, such as the NSF, the Army Corps of Engineers, the Navy, and the Department of the Interior, would continue to bear major fiscal responsibilities.

It is difficult to estimate the proportion of the proposed budget for NOAA that could be associated with physical research. In the report a breakdown for incremental cost over the decade 1971–1980 ($8.0 billion) is given. About 23 percent of this amount is allocated to research and education, and 21 percent to national projects. Both items are likely to be concerned with physical oceanography. A rough estimate is that between $100 million and $300 million per year of NOAA's incremental expenditure (that is, in addition to that which is now being spent) could be related in some way to physical research in the oceans.

The response of the federal government to this proposal has been the creation of NOAA as a subdivision of the Department of Commerce, rather than as an independent agency, with a 1971 budget of $310 million, rather than the recommended figure.

The largest element of NOAA is the Environmental Science Services Administration, which was formed in 1965 from the Weather Bureau, the Coast and Geodetic Survey, and the Central Radio Propagation Laboratory of the National Bureau of Standards. Also included in NOAA were a number of oceanographic programs from the Department of Transportation, the NSF, the Bureaus of Commercial Fisheries and Sport Fisheries and Wildlife, the Corps of Engineers, and the Navy.

There are some notable omissions from NOAA, as compared to the proposals of *Our Nation and the Sea*. NCAR remains in the NSF, and responsibility for IDOE has been assigned to this agency rather than

NOAA. Furthermore, NSF and the Navy continue to be the major sources of funds for four large oceanographic laboratories, which have some characteristics of national centers (Woods Hole, Scripps, Lamont-Doherty, and Miami). The original report had proposed that these laboratories be funded by NOAA.

It is too early to comment on the success of this reorganization. According to the data presented in Table IX.12, there has been no positive response by the federal government from 1966 to 1970 to the views expressed by successive Presidents, Congresses, and the news media that research in oceanography should be expanded. The data from the NSF program, which appear in Table IX.18, reflect the same trend. If the special program for ship support, now completed, is removed, funds for the oceanography program have decreased between 1966 and 1970.

Patterns of expenditure in oceanographic research differ greatly from both the atmospheric sciences and the geological sciences. According to Table IX.5, the 1968 manpower in each of these fields was 730 (oceanography), 5232 (atmospheric sciences), and 17,198 (earth sciences). However, according to Figure IX.14, research expenditures (excluding NASA, which has programs in these three fields) for 1970 are \$87.0 million (oceanography), \$118.1 million (atmospheric sciences), and \$78.3 million (geological sciences). The cost of ship operations presumably is responsible for the disproportionate costs per man of oceanographic research. Also, there is a relatively large military involvement; for example, in 1969, the Navy spent \$58.5 million on basic and applied oceanographic research. The applied research included some expensive items such as the Alvin program at Woods Hole.

TABLE IX.18 The NSF Oceanography Program, 1966-1971 (\$millions)[a]

	Program Allocations						
Fiscal Year	Biological	Geological	Physical	Marine Biology	Total	Ship Support	Total
1966	6.56	–	3.70	–	10.26	6.97	17.23
1967	6.85	1.94	2.97	–	11.76	6.38	18.14
1968	6.05	2.91	1.94	–	10.90	6.88	17.78
1969	3.39	2.55	2.16	2.90	11.00	8.64	19.64
1970	3.67	3.18	2.07	–	8.92	–	8.92
1971[b]							10.00

[a] In addition, the International Decade of Ocean Exploration was funded at \$15 million for fiscal year 1971. The Office of Polar Programs supports two ships (the *Eltanin* and the *Hero*), both of which are partly engaged in oceanographic measurements.

[b] Estimated.

Oceanography also differs from the other earth sciences in that contact between physical and life sciences is particularly close, with the life science component exceeding the physical science component in both number of scientists involved and amount of research activity. The field has had a curious recent history in terms of the enthusiasm generated among students and young scientists and the attention received from news media and from congressional hearings and governmental commissions. University faculty members are made aware of a very strong interest in oceanography among undergraduate and graduate students. The impression exists that oceanography offers large and rapidly expanding opportunities for young scientists. In fact, however, it is the smallest of the earth sciences in numbers of scientists and support, which has been constant or decreasing from 1965 to 1970.

The danger of an oversupply of scientists is obvious and may be reflected in AGI data for "other earth scientists" (in *Manpower Supply and Demand in Earth Science*), which include oceanography in an unknown proportion. This category shows that new employment will be fairly level, between 4000 and 5000 per year for 1970–1975, while degrees of all kinds will increase in number from 10,000 to 15,000 per year. Only in the category of PhD degrees does the forecast of production fall slightly below new employment in 1975.

8.4.4 *Space Sciences*

A degree of familiarity with NASA programs is necessary to understand in context the large expenditures in earth and planetary sciences. Three offices are directly involved: Manned Space Flight (OMSF), Space Sciences and Applications (OSSA),* and University Affairs (OUA). The OMSF budgets include manned spacecraft and associated boosters and development programs. These costs do not usually appear in science budgets, at least not in the data submitted to the NSF. However, a fraction of OMSF expenditure is assigned directly to science experiments and operations; an indirect charge to science also can appear through the ubiquitous budget category, "other," which includes, for example, Civil Service salaries.

The OUA made a substantial contribution to education in space sciences. Since 1965, however, the university program has decreased, and it will be phased out in fiscal year 1971.

* Since this report was written, OSSA has been split into two offices, one for space science and one for applications. The probable intention is to focus more effort on applications.

In this discussion, we are chiefly concerned with OSSA, whose programs include astronomy, communications, and life sciences, in addition to those of more direct interest to earth and planetary physicists, such as lunar and planetary programs, space physics, and parts of the environmental sciences program.

An important factor in assessing the cost of NASA programs is the cost of launch vehicles and "other" costs that include tracking, data acquisition, and administrative operations such as Civil Service salaries and costs of Center operations. These are classified as indirect costs and are not necessarily included in tabulated data. Between 1967 and 1972 (projected), launch vehicles accounted for 21 to 25 percent of the OSSA program, not including "other" costs. "Other" costs attributed to physics and astronomy programs accounted for 50 percent and 39 percent, respectively, of the total program costs in these fields. Table IX.19 presents a model of a "typical" OSSA program, and Table IX.20 shows the breakdown of the OSSA budget from 1967 to 1972 (projected). The proportion assigned to lunar and planetary science, space physics, and earth environmental sciences varies from 50 percent to 57 percent.

NASA works to a considerable degree in terms of specific missions, with tight time schedules, involving huge expenditures in certain peak years. Setting up such missions is a complicated process involving the views of Congress and the executive branch, the scientific and technical opinion of NASA headquarters staff and staff members of appropriate research centers, the preparedness of industry, and, last, the views of the science community, for science is the reason for the OSSA program, and outside scientific capability must be harnessed to achieve a satisfactory result.

The success of the collaboration between NASA and the scientific profession has varied greatly according to whether initiation or execution of a mission is at issue. With regard to execution, the collaboration has been fruitful and the achievements spectacular, but there are differences

TABLE IX.19 A "Typical" OSSA Science Program

Item	Cost (% of Total Program)
Science, spacecraft, etc.	40–45
Launch vehicles	10–15
Tracking, data acquisition, and administrative operations	40–45

TABLE IX.20 OSSA Annual Budget, Percentage Distribution by Program Categories ("Other" Costs Not Included)

Program	Percent per Fiscal Year					
	1967	1968	1969	1970	1971	1972 (projected)
Lunar	20	8	2	2	1	<1
Planetary	13	18	18	27	24	37[a]
Astronomy[b]	11	15	18	15	14	11[a]
Space physics	12	12	11	7	8	5
Life sciences (biology)	7	7	8	4	2	2
Space applications	12	18	22	24	29	24
Communications[c]	(6)	(6)	(7)	(8)	(7)	(10)
Earth environmental sciences[c]	(6)	(12)	(13)	(16)	(22)	(14)
Launch vehicles	25	22	21	21	22	21
	100	100	100	100	100	100
Total cost ($millions)	$576	$553	$453	$520	$566	$880

[a] Includes Viking, Outer Planets, Planetary Explorer, HEAO-A and -B, Phase B for HEAO-C and -D, ground-based astronomy, Large Space Telescope Phase B study.

[b] Does not include Apollo telescope mount.

[c] Not included in percentage calculation.

of opinion about the role of the scientific community in the initiation of space missions.

The science advisory structure is divided into in-house experiment selection and advisory groups; the NAS Space Science Board (SSB); and the President's Science Advisory Committee (PSAC). Key committees in the advisory structure are the Space Science and Applications Steering Committee (SSASC) and the Space Programs Advisory Council (SPAC). The SSASC reports to the Program Associate Administrator for OSSA and helps him to determine OSSA policy. It consists exclusively of NASA employees, but it has an infrastructure of subcommittees in astronomy, biosciences, ionospheres and radio physics, particles and fields, planetary atmospheres, planetology, and solar physics. These subcommittees are chaired by NASA program chiefs and typically draw half their membership from outside NASA. The main task of the subcommittees has been selection of experimenters for established missions; this has been one of NASA's most successful collaborative activities. Subcommittees have rarely been asked to discuss policy questions, however, and the SSASC does not necessarily respond to advice unless specifically requested. Nevertheless, these subcommittees have enabled a group of outside scientists to gain some understanding of the extraordinary complexities of the NASA organization and decision-making machinery. It is only with such experience that outside scientists can offer effective advice to NASA through any of the available channels. Therefore, it is unfortunate that a trend toward a few missions embracing a range of disciplines has led the SSASC to abandon standing subcommittees, replacing them with *ad hoc* subcommittees.

Five years ago, NASA created the Lunar and Planetary Missions Board and the Astronomy Missions Board to advise on future missions. A Physics Committee, with functions somewhat parallel to those of the Missions Boards, had a less important role in mission planning.

NASA appears to have been disappointed in the functioning of the Missions Boards, partly because they did not include a sufficient number of NASA employees (they were chaired by non-NASA scientists) and partly because their advice appeared, at times, to be inconsistent. Consequently, when Congress reduced the use of consultants in the 1971 budget, the Missions Boards were dissolved and replaced by SPAC, which is chaired by a non-NASA scientist, with NASA employees constituting no more than 25 percent of the membership. There are four advisory committees, on applications, physical sciences, life sciences, and space systems.

The outside advisory committees are PSAC, reporting to the President, and the NAS boards, in particular the SSB,* whose principal contact is with the Associate Administrator for Long-Range Planning and the Program Associate Administrator. The most important function of these bodies is the preparation of reports on different aspects of the NASA operation, such as space physics, lunar and planetary research, and space applications.

Space physics includes aeronomy; energetic particles, magnetic fields, and the interplanetary plasma; and relativity investigations. Aeronomy has already been discussed in Section 8.4.1. It was the only subject of space research in the 1950's before the advent of satellites. During the 1960's, after the discovery of the Van Allen belts, the interplanetary plasma became the foremost topic in the U.S. space program. This pre-eminence has now been so eroded that space physics as a whole is in jeopardy. Although relativity is frequently discussed, it has yet to be accorded special missions.

The status of space physics and its future requirements were reviewed in the following three reports:

Space Science Board, *Space Research: Directions for the Future,* Publ.

* While this report was in its final stages, the NAS and NAE responded to the reorganization of OSSA by creating an NAE Space Applications Board.

1403 (National Academy of Sciences–National Research Council, Washington, D.C., 1966).

Space Science Board, *Physics of the Earth in Space. A Program of Research: 1968–1975* (National Academy of Sciences–National Research Council, Washington, D.C., 1968).

Space Science Board, *Sounding Rockets: Their Role in Space Research* (National Academy of Sciences–National Research Council, Washington, D.C., 1969).

The second of these reports states the goals of the science community at a time when NASA was decreasing its activity in this subject area and planning still further reductions. The following excerpt from the report summarizes its purpose and content:

> The study defines a program of satellite, space probe, and sounding rockets missions for a concentrated attack on questions of fundamental physical mechanisms of the Sun–Earth system, in contrast with the exploratory survey that characterized the past decade. We place particular emphasis on coordinated investigations and on the development and utilization of new experimental techniques. We also stress the importance of organizing a major observational effort during the 1974–1975 period of low solar activity.

There follows a recommended program involving relatively low-cost satellites, sounding rockets, solar observatories, and "piggyback" instrumentation on planetary probes, particularly to the outer solar system. General recommendations concern the coordination of balloon, aircraft, and ground-based observations, the development of instrumentation, the need for coordinated research, the problems of data handling and analysis, and the need for NASA predoctoral traineeships to encourage growth of this important interface between physics and space sciences. The study does not deal specifically with the problem of priorities.

The fortunes of the NASA space physics program are illustrated by a letter dated May 21, 1970, from Homer Newell, NASA Associate Administrator, to the Physics Survey Committee. He points out that space physics was one of the strongest and most successful scientific programs in NASA but that it now has to be deliberately de-emphasized to encourage other programs. Even opportunities to fly small, inexpensive packages on Orbiting Geophysical Observatories (OGO) or Interplanetary Monitoring Platforms (IMP) are now few. The letter requested the Physics Survey Committee to treat this subject explicitly.

The funding history of space physics, projected to 1976, is depicted in Figure IX.19. The fiscal background for Newell's statement is forcefully illustrated. From 1970 on, the substantial elements in the program are

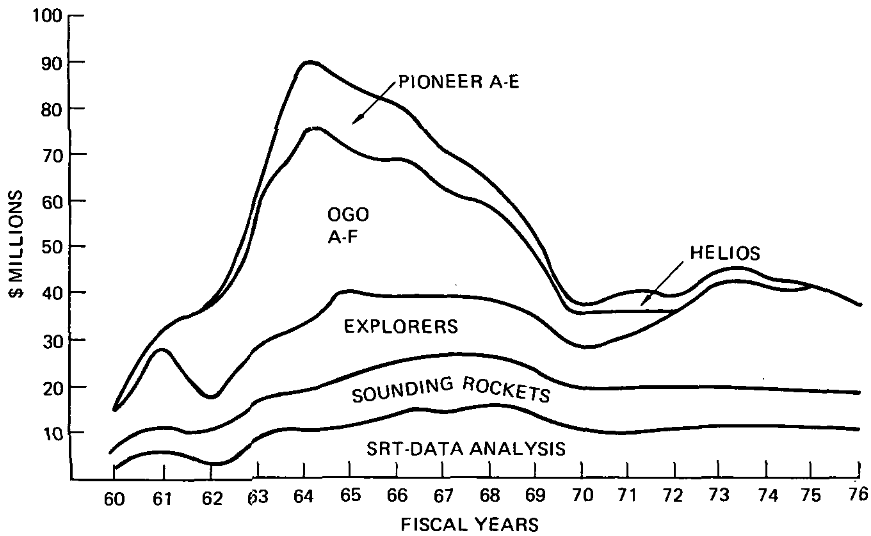


FIGURE IX.19 NASA support for space physics. The costs of launch vehicles, data acquisition, and administration operations are not included (see Table IX.19), nor is the cost of missions not flown principally for space physics.

rocket research, data analysis, and the variable-orbit Earth Explorer series. Of these three items, the rocket research and data analysis are continually in jeopardy because they do not represent a long-term commitment isolated from the rest of the NASA budget. They represent a flexible small program, highly valued by the scientific community but always vulnerable in NASA's continually changing fiscal situation.

Lunar and planetary research is largely dependent on NASA funding, although a small amount of ground-based work is sponsored by other agencies. The early years of NASA operation were characterized by extraordinary efforts to solve the engineering problems of space flight—not only for the manned programs but also for unmanned vehicles: the earth satellites, the IMP's, Surveyors, Explorers, and Mariners. During this period, engineering had overriding priority; science was supported when possible but was rarely the main purpose of a mission.

By 1965, it was clear that Apollo would succeed, that the space program would continue to appeal to Congress only if new programs replaced Apollo, that the engineering problems of unmanned flight were essentially solved, and that unmanned missions could be justified only on scientific grounds.

It was in this context that *Space Research: Directions for the Future* was prepared. In retrospect, this study had less influence on subsequent programs than was anticipated. Two problems were that the successor to Apollo had not been determined and that a successful working partnership between NASA management and the science community had not developed.

A number of studies have been issued since 1965: *Planetary Astronomy: An Appraisal of Ground-Based Opportunities*[26] deals with ground-based measurements; *Planetary Exploration 1968–1975*[27] discusses primarily the inner planets; *Lunar Exploration: Strategy for Research 1969–1975*[28] is concerned with lunar research; *The Outer Solar System*[29] treats the outer planets; and two general studies that include lunar and planetary exploration as essential parts of the overall space program, *The Next Decade in Space*[30] and *A Long-Range Program in Space Astronomy.*[31]

These reports have the same general theme. Great enthusiasm exists in the science community to explore the solar system and thus to learn more about its origin, the origin of life, and the processes on this planet. Unmanned missions are capable of performing the needed investigations, although the unique opportunities of Apollo should be used. Priorities can be established; however, continuous consultation between scientists and management is necessary to ensure that these priorities remain valid as new discoveries change the picture.

Current NASA plans for lunar and planetary exploration are ambitious. In addition to present plans for a Venus–Mercury flyby, a Mars Mariner orbiter, and the Viking Mars lander, there also are plans for Grand Tours* to the outer planets and low-cost Planetary Explorers to Venus. A Mariner Comet mission also has been suggested. Planned support for planetary investigations increases from $154 million in 1971 to $461 million in 1973. In addition, lunar exploration, which does not include the costs of manned space flights, is budgeted at $8 million in 1971 and is scheduled to increase to $21 million by 1975.

We will briefly examine the relationship between NASA and the scientific community with respect to lunar programs. Apollo could not be justified in purely scientific terms; it was a national adventure in the spirit of the journeys to the Poles and the ascent of Everest. Once the goal was achieved, there was no longer a strong appeal to the imagination of the general public or, in particular, of members of Congress. The inherent dangers of the mission and the need to develop a successor† in the field of

* While this report was in preparation, Grand Tour was canceled in favor of a Mariner Jupiter–Saturn mission.

† The Space Shuttle has now emerged as the Apollo successor. The repercussions for space science will be great if this program is funded, but it is too early to predict the total effect.

manned spaceflight have become prime considerations in NASA as the Apollo program is curtailed and terminated. At the same time, analysis of lunar samples has rapidly increased scientists' knowledge of the origin and history of the moon–earth system, and revolutionary new data are anticipated from every future Apollo flight. The scientific community has been insistent in its demands that Apollo be flown out, but NASA has met only part of these demands.

This conflict between management and scientific interests dominates the present relationship between the lunar scientific community and NASA. As the Soviet Union demonstrates the feasibility of unmanned sample returns, the question often raised is whether an unmanned program could have avoided some of the conflicts of purpose and thus have produced a healthier long-range program.

The NAS study *Space Research: Directions for the Future* was undertaken against the background of the euphoria created by maximum NASA budgets and the imminent success of Apollo. At that time, the planetary program favored by NASA was the Voyager lander. This mission, to cost about $2 billion, was to place a complex biological laboratory on the surface of Mars to investigate the possible existence of life on that planet. The mission was never fully defined, and, despite a great deal of work, it proved too ambitious for the existing capability. It would be useful to study this project and the lessons it provides in regard to the complexities of management decisions and relationships with Congress and the scientific community, for it is increasingly clear that the strength of NASA lies in its ability to implement a program, once a strong national commitment has emerged, but that the more subtle processes needed to develop a program with satisfactory long-term support are less well understood.

Planetary Exploration 1968–1975 was written after the cancellation of Voyager. The scientific objectives were the same as those in the 1965 study: to understand the origin and evolution of the solar system, the origin and evolution of life, and the dynamic processes that shape man's terrestrial environment. The principal general recommendations of this study were, first, that the planetary exploration program should be presented as a contribution that exploration could make to a broad range of scientific disciplines rather than as a single goal, and second, that the use of small, relatively inexpensive spacecraft should be an essential element of the planetary program. Pioneers were suggested for the outer planets, and what is now called the Planetary Explorer, for the inner planets.

The mission recommendations of this study were in order:

1. Small spinning spacecraft to all planets at all opportunities (Pioneer, Explorer)

2. A 1971 Mars orbiter followed by a Mariner-class Mars lander for biological purposes
3. Venus–Mercury flyby
4. Multiprobe mission to Venus
5. A major lander on Mars

The significance of the words "Mariner-class" (in 2) and "major lander" (in 5) for Mars is important. The intention was to give lower priority to an expensive, Voyager-type mission. Cost estimates for the Mariner-class mission were $350 million.

Venus: Strategy for Exploration[32] deals with developments in the Planetary Explorer; this report shows that orbiters, probes, and landers for Venus are all compatible with small spinning spacecraft of minimum cost. Thus, item 4 is shown to be part of item 1. The report repeated the underlying reasons for the science community's continued assignment of higher priority to many minimum-cost programs than to single, large projects. The Venus–Mercury flyby is approved and funded. Pioneers F/G for Jupiter are also funded. The Mars 1971 orbiters are launched. Funds for the Viking lander have been provided for 1975. This program is a halfway step between the recommended Mariner program and the low-priority Voyager; it is currently estimated to cost $750 million, which has created a funding problem in the planetary program.

This funding pattern was made more difficult by the addition of a Grand Tour program using TOPS spacecraft, from which NASA had to back off to Jupiter–Saturn missions based on Mariner spacecraft.

The considerations behind this change of program are contained in two NAS reports, *The Outer Solar System*[29] and *Outer Planets Exploration, 1972–1985.*[33] Two issues are involved. The first is the general need to explore the outer solar system, beyond Jupiter and Saturn. To do so requires new long-life technology and, for direct flights, low-thrust ion propulsion. To avoid ion thrust, use must be made of the gravitational boost available at Jupiter. Saturn is frequently accessible by this means, but reaching Uranus and beyond depends upon planetary concatenations that are rare.

Much emphasis has been placed on the fact that the planetary lineup in the 1970's and 1980's occurs only once in 180 years and permits us to visit three or four planets with one spacecraft. This is the Grand Tour concept, and it is the second major issue included in the reports under discussion. In fact, the extreme rarity of the five-planet lineup may not be the most important issue: the fact that even Uranus may not be accessible again until the next millenium is of great concern to some planetary physicists.

The Outer Solar System and *Outer Planets Exploration, 1972–1985* both recognize a duality in strategy for the outer planets. The first point of view emphasizes exploration of as much of the planetary system as possible, even if quantitative results are difficult to obtain. This point of view favors the Grand Tour concept, the TOPS spacecraft, and imaging techniques. The alternative view stresses the diagnostic approach and a desire for investigations to answer specific questions about the physical processes taking place in the planets. For this, probes and orbiters are the ideal tools, and Jupiter and Saturn are the prime targets. Long-life technology and Jupiter gravitational assistance are not essential and lower-cost spacecraft, such as Pioneer or Mariner, are useful vehicles. The change to Mariner Jupiter–Saturn missions thus emphasizes the diagnostic approach.

NASA has, in general, followed the recommendations of the NAS in the planetary program. The most significant differences are that Pioneer or Explorer missions to the inner planets (specifically to Venus) have not been funded, while Viking has been funded at a higher level than was proposed. In addition, Grand Tour was brought in at a rather high cost, which some believed could ultimately have been more than Viking (Grand Tour included at least four missions, in contrast to Viking, but the budgetary problem has similarities).

These differences, though they may not appear to be large, probably represent an important distinction between the approach of space scientists and the needs of NASA as seen by its top managers.

The science community, in virtually every report, draws attention to the value of many opportunities at minimum cost, whether they be Explorers, Pioneers, rockets, or even balloons and ground-based observations. From the scientists' point of view, flexibility of missions, involvement of the maximum number of people (including graduate students), a steady level of effort, a sequential learning process—all characteristics of small programs—have great importance in the development of a successful and healthy space science program.

From the management point of view, the large, single-opportunity programs have real advantages: The management problems are easier than for many small programs, and the large programs with long lead times entail a commitment from Congress and the Executive Branch for a substantial period (from five to ten years). The organization of NASA shows to its best advantage in relation to large commitments, particularly if new engineering concepts are involved. The large program has peak funding years that produce a new level of effort in the program that, hopefully, can be maintained. It is easier to defend funding for a single large project than a continuing program whose flexibility makes it a tempting target for delays or economies. Finally, NASA management believes that it is easier in practice to convince Congress and the public of the merits of programs such as

Viking and the Grand Tour than to convince them of the value of smaller programs, which may appear to have a flavor of self-indulgence by a favored profession and whose multiple objectives are hard to explain in easily understandable forms.

A compromise between these two positions might be reached more easily if they were not to some degree mutually exclusive. A large commitment tends to place emphasis on successful execution and to foreclose alternative scientific options. The science community likes to keep its options open, while good management practice requires a final commitment to a particular experimental plan, a commitment not easily obtained from a diverse group of individuals who thrive on new ideas.

When projects with large peak funding are introduced, the total support rarely responds completely to the peak needs; sometimes, it does not respond at all. Since the program is committed at this stage, there remains no alternative but to prune other existing programs. Some programs benefit from this pruning, but there is also a pressure to economize on small, flexible programs, simply because they are flexible and can be turned on and off on short notice. Whether this analysis is correct is less important than the fact that many believe it to be; thus, it is a source of concern to the science community.

The space applications section of OSSA is only partially related to earth and planetary physics. Communications, agriculture, ecology, cartography, urban studies, and the like belong to other disciplines. However, the general concept of Earth Observation Satellites can be used for a variety of purposes, including meteorology, oceanography, and geophysics. Reports addressed specifically to NASA's role in such research are

The Terrestrial Environment: Solid-Earth and Ocean Physics[24]
Useful Application of Earth-Oriented Satellites[34]
Remote Sensing with Special Reference to Agriculture and Forestry[35]
Plan for U.S. Participation in the Global Atmospheric Research Program[16]

The applications program differs from all other OSSA programs in that NASA does not consider itself responsible for all its aspects. NASA's role is that of an entrepreneur, with special skills and insights into space techniques, not available to other agencies but important to the missions of those agencies. NASA attempts to stimulate interest in relevant agencies and to demonstrate the feasibility of certain observational techniques. Actual operations and the justification for the operational phase are the responsibilities of the user agency.

This procedure has been particularly successful in meteorology and, by means of harmonious cooperation between NASA and NOAA, has led to the inception of GARP and WWW. Other aspects of the applications program are less secure, and difficulties lie ahead for user agencies. The initial stages in an applications program are characterized by an unfamiliar surfeit as NASA applies its vast resources to an area in which the rate of advance may have been slow. The next stage of development requires reorganization and changing attitudes in the user agency. Such an experience can be both difficult and rewarding; however, the effort must be made if the resources of NASA are to be used without detracting from agency responsibilities.

Priorities for Space Research 1971–1980[11] is the most recent Space Science Board study of the NASA program. This study, lasting about nine months, was led by Herbert Friedman. A 14-member Executive Committee and discipline-oriented working groups, with a total of some 90 scientists, including foreign consultants, participated in the most intensive phase of the study, a period of three weeks. The study represented a departure from previous Space Science Board studies in that it produced a complete ordering of priorities. It did so by considering fixed budget levels and approximate costs of major missions.

To achieve this objective, the study was restricted to the OSSA budget. This was recognized as a serious limitation, but the resources available to the study were inadequate for an examination in depth of more than one relatively homogeneous budget area. A broad study of the entire national space effort was visualized as a necessary subsequent step.

Three budget levels were considered: a base level corresponding to the existing level of effort (approximately $550 million per year; see Table IX.20); an intermediate level, 25 percent above the base level; and a higher level, 50 percent above the base. The base-level program allowed few new opportunities. The higher budget level would permit a broad-gauged and more balanced program of scientific exploration and applications. The intermediate budget level would allow a few new projects of high scientific merit and technological challenge.

It was considered to be unrealistic to develop firm conclusions that would apply ten years in the future. Consequently, the study called for a continuing process of priority assessment and recommended that NASA and the NAS consider how to achieve the most effective coupling between NASA's internal planning and the advice of the Space Science Board.

Other problems related to the means by which a priority judgment was made. The criteria for intrinsic and extrinsic merit developed by A. Weinberg were used, but many other factors also had to be considered. A number of these factors related to the peculiar characteristics of the space program; for example, are space systems the only or best approach to the problem? It is virtually impossible to apply a consistent set of criteria to subjects so different in their justification as galactic astronomy and geodetic surveys. The study also was influenced by such factors as the national interest in the quality of the environment.

Finally, since priorities were based on fiscal limitations, it follows that a low-cost program was favored over a high-cost program, because it required less sacrifice on the part of other disciplines. However, costs are difficult to assign if the mission has novel features, as is frequently the case. Cost overruns occur in all high-technology areas. NASA is not particularly at fault in this matter, but a factor of 2 on a high-cost item can be the determining factor in the eyes of a diversified scientific community. The Viking cost increase is an example. On a large mission, such as the Grand Tour, a large overrun could put an end to many other programs of equal or possibly higher scientific value.

The problem of dealing with programs having different costs in a single budget structure pervades this priority study. The Executive Committee stated:

> Large complex missions seriously imbalance the program unless a solid base of smaller missions is continued. The risk of such unbalancing is often increased by the dramatic appeal of the large missions to those more concerned with nonscientific justifications. From a management point of view, program choices are influenced by the desire to fit missions to special capabilities of NASA centers in an effort to preserve balance of effort in various centers. *Nevertheless, we feel strongly that such large missions should not be mounted if they crowd out the smaller missions of high scientific priority.*

(This subject is discussed in Chapter 5 and is the substance of one of three recommendations presented in that chapter.)

Table IX.21 summarizes the recommended new starts, that is, additions to existing programs. No new starts were recommended in space biology. Some important recommendations for completing Apollo were addressed to OMSF. The main feature of this listing of new starts is that a diverse group could agree unanimously on priorities among different fields. The emphasis on small programs is also clear. Thus, in the planetary program, Planetary Explorers to Venus and Pioneers to Jupiter were placed ahead of the Grand Tour. In the astronomy program, one expensive item (High Energy Astronomical Observatory) appears in the base program, but Small Astronomical Satellites, increased rocket, aircraft, and balloon opportunities, and mirror technology are placed before all other more costly missions. Solar-terrestrial physics emphasizes rockets, balloons, and data analysis before all satellites and probes, with one exception.

The Executive Committee did not agree in all respects with the con-

TABLE IX.21 Recommended New Starts in the OSSA Program at Different Budgetary Levels

	Budget Level		
Program	Base	Intermediate	High
Planetary exploration	1. Planetary Explorers to Venus 2. Pioneer Jupiter diagnostic missions.		1. Solar electric Mercury orbiter 2. Grand Tour (TOPS) 3. Interstellar/heliosphere particles and fields
Lunar	1. R and D on automated lunar landers		1. Lunar orbiter
Astronomy	1. High Energy Astronomical Observatory 2. Small Astronomical Satellites 3. Rockets, balloon, aircraft astronomy 4. Mirror technology for large telescope	1. 1.5-m space telescope 2. Orbiting solar observatories	1. Solar observatory with precise guidance 2. Kilometer-wave orbiting telescope
Gravity physics	1. Earth-orbiting and sun-orbiting satellites		
Solar-terrestrial physics	1. Coordinated satellite measurements 2. Rockets and balloons 3. Data analysis	1. Atmospheric explorers (aeronomy) 2. Three-dimensional solar-wind measurements	1. Explorer satellite for various plasma studies
Earth observations	1. Earth observatory satellites 2. Small applications technology satellite 3. Satellite-to-satellite tracking 4. Expanded aircraft surveys	1. Earth Resources Satellite 2. Synchronous earth observations satellite 3. Sea-surface measurements	1. Recoverable Earth Resources Satellite

clusions of the discipline-oriented working group papers. The Planetary Working Group, for example, endorsed exploratory-type Grand Tour missions. The Executive Committee, on the other hand, recommended a concentration on diagnostic missions to Jupiter. Their reason was the impossibility of accommodating the Grand Tour mission with other high-priority new starts in the base program.

8.5 DISCUSSION AND RECOMMENDATIONS

8.5.1 *Costs and Benefits in the Earth Sciences*

The earth sciences are primarily applied sciences, many aspects of which relate directly to human needs. A large directed or mission-oriented research component, justified in terms of potential economic gains, the saving of lives and property, improvements in the quality of life, and the like, necessarily exists. These considerations are reiterated in most NAS reports on the earth sciences, and the magnitude of potential benefits is often cited as the rationale for a particular level of effort.

This Panel does not argue that potential benefits should be used as the only, or even the main, justification for research and development costs, even in a largely applied area. The value of the unknown obviously cannot be assessed in terms of predictable benefits, and most major advances are based on unanticipated new discoveries. No acceptable criteria for the needed undirected research have been developed yet in the earth sciences; there exists only a communal commonsense judgment by groups of competent and motivated scientists working in the field.

However, these considerations should not obscure the fact that there are areas in the earth sciences in which the benefits can be estimated, the physical principles are reasonably clear, and a large effort, mainly technological, probably will produce immediate results.

In previous decades, the possible benefits from research, directed or undirected, in most of the earth sciences was orders of magnitude larger than probable costs. Industry and government did not have the working partnership and the modern technology that underlie the capability to mount huge efforts at short notice in almost any area related to the physical sciences. Such capability now exists. Its potentialities in the earth sciences are almost unlimited. But the cost can be high, not only in dollar expenditures but in manpower resources, for such efforts demand the attention of skilled scientists and engineers whose numbers cannot be readily increased.

Are costs approaching possible benefits in major projects in the earth sciences? The disturbing aspect of this question is that the effort devoted to answering it is small. A meteorological study quoted in this chapter gives benefit-to-cost ratios from 10 to 100 for meteorological research and operations. This estimate would be satisfactory if one could be sure of the conclusions. However, the study was a one-man effort, with access to published data only, and takes little account of the explosive increase in expenditure that could accompany a successful conclusion to GARP and WWW. It is also a human weakness to justify a personal commitment with an exaggeration of benefits and a shading of costs. Under these circumstances, a factor of 10 is not a comfortable margin to justify a very rapid expansion of effort in meteorological operations.

In geophysical research, constraints are exercised on expenditure because mineral exploration is in the hands of private enterprise. However, possible future projects, such as earthquake prediction and control, could be immensely expensive, and the benefits could be identified and evaluated.

Oceanography has an important military justification, and the interest of successive U.S. presidents in a new frontier of knowledge has stimulated support for it. As the subject matures, however, problems in benefit-to-cost ratios, parallel to those in meteorology, can be expected.

We do not wish to overemphasize the role of cost–benefit studies in the planning of science. However, there are programs in the earth sciences whose costs can be huge and for which a satisfactory benefit-to-cost ratio is an important consideration. In these instances, cost–benefit studies should be pursued with more energy and more depth than has previously been the case.

8.5.2 *Funding in Geophysics*

Funding in geophysics appears to have been particularly slow to follow the revolutionary advances of the past five years. No breakthrough of equal importance has occurred in the previous history of geophysics. It could be argued that such dramatic new ideas have never developed so rapidly in any of the earth sciences at any time in their long history. Funding in the NSF Geophysics Program has not increased during the recent inflationary period; therefore, the level of research activity must have decreased. In addition, other federal agencies are reducing their research commitment in fields not directly related to their missions, which has resulted in increasing demands on NSF funds.

This Panel has considered the reasons for this apparent failure to rise to a great new opportunity in geophysics and, specifically, the role of the NSF. Other agencies have been inhibited by the operation of the Mansfield Amendment, but the NSF is clearly interested in responding to the requirements of the science community in such an exceptional case. However, the community must present its case in satisfactory terms. As we discussed in Chapter 5, the creation of new programs or large increases in support for existing programs depend on a complex process that involves strong leadership by experienced and knowledgeable individuals, encouragement and assistance from an important section of the community, activity through NSF channels, and, finally, response by agencies, the Office of Management and Budget, and Congress.

The data presented in this chapter offer evidence that the NSF has responded reasonably well to the new opportunities in terms of national programs. The JOIDES program ($7.2 million in 1971) is solely for ocean-floor drilling and caters directly to the interest in sea-floor spreading. More than $5.0 million of IDOE expenditure in 1971 will also be for this purpose, as are important parts of the NSF Oceanography Program and the Polar Programs.

Funds for the NSF Geophysics Program, which principally supports individual research workers, have been stationary; therefore, capability has decreased in recent years. However, a comparison of Table IX.17 (for geophysics) and Table IX.14 (for the atmospheric sciences) shows that proposal pressure has been almost identical, including the unusual situation developing in the early part of 1971. In its internal budgetary process, the NSF uses proposal pressure as a means of comparison among disciplines. An increase in activity by the science profession should be registered by this means. On this basis, geophysics has prospered as well as was justified; certainly, it has fared no worse than have the atmospheric sciences.

We conclude that the responsibility for the present fiscal doldrums in geophysics lies more with the science community than with the NSF. The traditional independence of geophysicists and the relatively small size of their projects militates against communal enterprises of the kind that generate new activity and bring about higher levels of support.

This situation is likely to change. The Geodynamics Project, now being developed both nationally and internationally, is an attempt to effect a synthesis among small programs and to fill gaps, so that advance can occur evenly over a wide field. We believe that this project, when it has fully evolved, should receive enthusiastic support from relevant agencies, and we urge them to follow its development with a view to early implementation of its proposals.

8.5.3 *The Movement of Physicists into the Earth Sciences*

Table IX.7 shows the relatively large proportion of non-PhD's engaged in research in the earth sciences. This finding suggests the possibility of a general upgrading in the subfield and potential benefits from a greater intake of physics PhD's.

In recent years, physics departments have attracted a large proportion of the most able graduate and undergraduate entrants to universities. The earth and space sciences now have the opportunity to recruit physics graduates into their ranks. We have made appropriate recommendations in Chapter 4 on the educational aspects of this situation. There is also a short-term aspect related to a decrease of employment opportunities for

PhD's in the core subjects of physics and an increase of student interest in the earth and space sciences. Unfortunately, the earth sciences have been unable to react in a significant manner to this unusual and probably ephemeral opportunity. The NCAR postdoctoral fellowships program (illustrated in Table IX.8) is the most constructive response known to the Panel, but it is small in proportion to the needs. In the universities, it has become increasingly difficult to support postdoctoral research because of federal economies. In federal research laboratories, retrenchment under Civil Service regulations makes it particularly difficult to accommodate young newcomers to the field.

The Panel does not believe that a physics PhD is necessarily a valuable addition to earth-science research programs. He will lack experience, and, more important, he might not have the motivation and dedication needed to be an effective research worker in this field. However, a first-rate physicist who can become deeply involved should be a welcome recruit. Many of the young scientists currently being turned away from the field could have this potential.

We propose that steps be taken to encourage young physicists to enter the field, at least on a temporary, trial basis. The immediate postdoctoral period is the most important consideration. A temporary surplus of PhD's in the earth sciences could result, and the problem of accommodating active young scientists in permanent positions would have to be faced as an immediate consequence.

The time for effective action may be passing, but we urge the appropriate agencies to review the following possible courses of action:

1. To put immediate risk money into postdoctoral fellowships in universities, national centers, and federal laboratories specifically for transfers of young physicists into the earth sciences
2. To consider procedures, possibly involving variations of Civil Service rules, to absorb research workers in the "holding pattern" created by postdoctoral fellowships in the earth sciences into permanent employment
3. To accord high priority to those plans for expansion in the earth sciences that would profit from an intake of young physicists

With regard to the third suggestion, we note that no existing NAS studies in the earth sciences have dealt with this aspect; therefore, we are advancing a new and temporary basis for priorities. Research centers are the most obvious candidates for a higher priority ranking. However, the Panel does not believe that national centers should be the only model. For example, decentralization into regional research centers also should receive consideration.

8.5.4 *The Rationale for a Space Science Program*

Priorities for Space Research 1971–1980 discussed the OSSA budget at three fixed levels; however, it did not consider the rationale for a space science program in the overall framework of U.S. science. As a result, we have no adequate discussion of the reasons for expenditures on space projects, as compared, for example, with expenditures on earthquake prediction or atmospheric quality control.

Such questions probably will be asked insistently in the coming decade, and they can and should be answered. We believe that the case in favor of a space science program at an appropriate level of effort can be justified, although it involves a number of different questions that require careful analysis.

1. The origin of life and the origin of the solar system have an unusual fundamental significance. Such questions have stimulated man's imagination since earliest recorded time and are an intrinsic part of his culture.
2. Direct applications are an important part of the OSSA program. Subject to ultimate cost–benefit analysis, they obviously should be supported.
3. Indirect applications also exist. The assertion that increased knowledge of other planets aids in understanding our own can be supported by detailed arguments.
4. NASA supports large in-house activities. There is a point beyond which the dismantling of this capability makes little sense and could cause serious national problems.
5. New ideas in U.S. technology have been dramatically stimulated by the space program.
6. If federal programs are required to adapt rapidly in response to changing national priorities, the United States ultimately could be obliged to adopt manpower policies to mitigate hardship and to preserve national capabilities in fields involving highly skilled personnel. The NASA program offers a generally benign and nonprovocative tool for this purpose; in this perspective, the unfortunate feature of the Apollo program is that its start could not have been delayed until 1970 when military budgets were being reduced.

These and other aspects require careful assessment. The priorities study did not address them, nor did it address the complicated factor of manned space flight, except to state that in the post-Apollo era, man is not required for any of the recommended science or applications programs. This assertion may still be valid in the context of the space shuttle program if we distinguish between man as a pilot and navigator and man engaged in other activities. In the Apollo program, man's functions could not be separated in this way, but it is now desirable to understand the rationale for manned spaceflight *per se,* the independent rationale for space science, and the extent to which the two can interact and support each other. *We recommend to the Physics Survey Committee that the National Academy of Sciences and the National Academy of Engineering institute a study of the rationale for a space science program in the broad context of U.S. science and technology.*

8.5.5 *The Future of Space Physics*

The main components of the OSSA Space Physics Program are aeronomy and particles and fields. Both subjects had periods of rapid buildup—aeronomy because it represented the scientific frontier before satellites and space probes, and particles and fields because the discovery of the Van Allen belts was the first major breakthrough in the U.S. space program after the first Sputnik. Both went through a period in which NASA funding and facilities were readily available. For example, the Mariner space probes were specifically designed to investigate the planets, yet an important part of their capability was always assigned to interplanetary plasma work.

The profession responded to this stimulus. The Space Physics Section of the *Journal of Geophysical Research* publishes 10 to 12 issues per year and contains articles almost entirely devoted to aeronomy and particles and fields. A typical issue, picked at random (Volume 76, No. 10, April 1, 1971), contains 25 major articles and 6 letters covering 320 pages.

This research momentum persists while NASA reduces expenditures in space physics in order to invest in other areas. This problem was discussed in the second part of this chapter, and the specific request of the Associate Administrator for Long-Range Planning for a reaction from the Physics Survey Committee was mentioned.

The view of the science community has been expressed in the two parts of *Physics of the Earth in Space* and in *Priorities for Space Research 1971–1980.* These documents tacitly recognize that space physics no longer has the primacy that it once enjoyed. Nevertheless, they protest strongly against the rapidity of deflation of funding in this area. They ask for increasing support for items of relatively low cost, such as rockets and balloons, and increased funding for data analysis. They also ask for specific missions that will enable the subject to advance beyond its exploratory phase.

The proposed OSSA space physics budget given in Figure IX.19 shows a partial response to these views. From 1970 on, level funding is proposed, with a slight increase in support for data analysis.

The views of the space physics community might appear to be special pleading and might be discounted for that reason. However, it should be noted that the modest new starts in space physics proposed in *Priorities for Space Research 1971–1980* were supported by all space science disciplines in the knowledge that these efforts would compete with their own projects. The recommendations of the report reflect a feeling of concern of the whole science community about the funding history of this field.

Scientific enterprise must be encouraged if the United States is to continue at the frontiers of research. It is mandatory that a subject receive increasing support when new ideas are developing and new discoveries are imminent. When funds are limited, less active areas should lose support to make way for more exciting developments.

Concern about the space physics program was voiced in this context; it implies that both the rate of buildup and the subsequent rate of decrease of support were too rapid for the health of the subject. The changes have been too sudden to be accommodated in the structure of the space sciences at universities and have resulted in a loss of morale and a lack of confidence in the intentions of supporting agencies.

It is difficult to offer useful guidance in this problem of scientific management. However, we feel that the support recommended for space physics in the report on priorities for space research must be regarded as minimal. We also recommend that NASA management adopt a more moderate role in regard to programmatic changes. It is appropriate for Congress and the public to call for changes in federal programs, but, when implementing such changes, NASA management can regulate the pace of change according to its perception of the health of the scientific community affected. NASA rightly considers the nature and composition of its in-house research staff and facilities when reprogramming; similar solicitude for the university and industrial science community, especially when it has expanded its efforts in response to NASA incentives, should be exercised.

8.5.6 *NASA Earth Observations Program*

In his letter to the Physics Survey Committee, Dr. Newell asked, with respect to the Earth Observations Program, "Could your Committee define better the role to be played by the physics community in solving these problems and suggest ways by which NASA can assure more intimate involvement of the most capable people?"

NASA has adopted the role of an entrepreneur in the Earth Observations Program, leaving it to user agencies to define and justify scientific goals and to take over the support once feasibility is demonstrated. We regard this attitude as correct and effective, at least as far as meteorological developments have been concerned, and probably also in other fields. Therefore, it is not the responsibility of NASA to plan the eventual manpower requirements in space applications. However, we.believe that there are interface areas in which NASA should deploy more effort, probably by reprogramming existing manpower rather than by recruiting new personnel.

NASA has done well in providing systems and instruments broadly directed toward applied needs. Some of the user agencies have reacted by modifying their approach to take account of the new methods available. There remains the most difficult task of optimizing the space vehicle, the instrumentation, and the needs of the user agency to obtain the best result. In this area, the split responsibility between NASA and the user agencies could have its most serious disadvantages and must be handled with the greatest care.

The GARP program offers an example of this problem. The precision of the data available from satellite information is, even under ideal circumstances, barely adequate for the purpose of prediction on the time scale of ten days. Choices must be made among different infrared sounding systems; the spatial resolution and the spectral regions, the spectral resolution, the signal-to-noise ratio, all can be varied. Microwave sounders can be used. Occultation experiments can yield additional valuable data. Winds can be measured at the sea surface from constant-level balloons and from cloud motions. Information from cloud morphology can be used in a wide variety of ways.

The best choice of observations must be matched to the numerical computation, which can be handled in many ways; these operations require a communication and data distribution network that also can impose limitations on the system. There are no margins available. Ultimately, an optimization of the entire system, consisting of observations, distribution, and numerical computation, will be required. The interface problems entailed by this system optimization are likely to be neglected when there is dual agency responsibility. Some important work on this optimization problem is already taking place at the Goddard Space Flight Center Institute of Space Studies. This is probably the most significant effort in the field, but it is small in comparison to the need. An expansion in this area would be a most valuable step for NASA to take in support of the Earth Observations Program.

The meteorological component of the Earth Observations Program is well defined compared to the Earth Resources area. Imaging techniques exist that could be used for many different purposes, of importance principally to ecology. This area has not involved physics or associated disciplines in the past, but it is one in which physical research methods could be usefully applied. Major problems pertaining to the Earth Resources Program are development of equipment, interpretation of its output, handling of data, and application to significant problems. These problems also tend to fall in the interface between agency responsibilities. The situation has analogies to that which gave rise to the concept of operational research during World War II. Physicists made major contributions to this field. A relatively undirected effort, involving physicists and members of other disciplines, with the objective of identifying the problems and means of solution is appropriate if the Earth Resources Program is to maintain its momentum. Such an effort could be sponsored or undertaken by NASA. At the present time, it should not be difficult to find enthusiastic recruits from the physics community.

In a recent review[36] of *Priorities for Space Research 1971–1980,* the Washington correspondent for *Nature* interprets the study and NASA's response in terms of a major confrontation between this agency and the science community. Most of the scientists concerned would not view the situation in such apocalyptic terms; nevertheless, there is a certain feeling of disenchantment with the advisory process, particularly in regard to the planning of new missions.

There are dangers in this situation. NASA will be subjected to critical scrutiny by all branches of the government during the next few years. The manned program will probably be de-emphasized. If the military component does not become of overriding importance,* space science and applications could become the main justification for a continuing program, and a science rationale supported by a substantial proportion of the profession will then be essential.

The early years of NASA were characterized by a specific mission, Apollo, and the solution of engineering problems of space flight. The science community was involved in the secondary capacity of recommending the optimum use of a given capability; discussion of the nature of the missions was minimal. The disciplinary subcommittees were an important interface between scientists and the agency, and they operated remarkably well.

* Recent statements on the Space Shuttle appear to envisage approximately equal use for military and nonmilitary purposes. Virtually all of the nonmilitary use is for space science and applications; medical and biological studies may account for a significant portion of the effort. This section was written before the shuttle was proposed to Congress, but the conclusions are still valid.

The engineering problems of spaceflight have on the whole been solved: At least these considerations no longer dominate the decision-making process as they did in the past. The question of the optimum mission from the point of view of science and applications has become increasingly important. To meet the need for advice in the area of mission planning, NASA formed the Missions Boards; presumably the agency was disappointed in their performance, for they were abolished as were also the standing disciplinary subcommittees.

The Space Programs Advisory Committee (SPAC) is a new organization whose effectiveness has yet to be tested. The disciplinary subcommittees have been replaced by *ad hoc* groups whose membership will continually change. Therefore, SPAC and the NAS-NRC Space Science Board constitute the formal interface between NASA and the science community. These two bodies have an important role to play in the formulation of future missions and other NASA policy. Nevertheless, we believe that the interface is too narrow, in regard to both the number of scientists involved and the time spent in their deliberations, to fulfill all the functions needed to bring the scientific community and the agency into a working partnership. Scientific advice should be formulated at many different levels. It is not possible, for example, to give effective advice on scientific priorities for missions without knowledge of costs and feasibility and the many other factors involved in formulating policy. It is lack of experience by the scientific community on these matters that sometimes contributes to an impression that their best advice is not followed.

Therefore, in the opinion of this Panel, a more extensive advisory structure to supplement boards such as SPAC and the Space Science Board is desirable.

Another important factor in the communication between NASA management and the scientific community relates to the absence of active research workers in Washington with direct access to the decision-making process. In the subtle balancing of scientific, engineering, managerial, and political factors, a scientist communicates best and with greatest confidence with his research peers. This was an important consideration in locating the Goddard Space Flight Center in the Washington area. However, the distance to the Maryland suburbs seems to be an impediment to close communication, and Goddard's function does not differ greatly from that of other centers with substantial scientific efforts.

At one time, NASA considered the appointment of a Chief Scientist to work at headquarters. We believe that the appointment of a scientist of distinction to this position could greatly improve the relationship between NASA and the scientific community. The Chief Scientist should have the opportunity to create a small institute, possibly similar to the Institute of Space Studies in New York, where active research can take place. He should be without program responsibility, and his main function should be to represent the views of the science community directly to the Administrator and to key decision-making groups.

A device employed by the NSF to increase communication between program staff and the science community is to obtain the temporary services of university scientists on leave for periods of a year or longer. This procedure also would be helpful to NASA, and more vigorous attempts should be made to institute such a policy.

To bring about a closer relationship between NASA management and active research scientists we recommend that the Physics Survey Committee consider the following measures:

1. Greater reliance should be placed on standing disciplinary subcommittees in which experience can be gained over many years. We suggest that these subcommittees be used more frequently as an informal forum for discussion of future policy.

2. There should be science planning groups, composed of representatives of both NASA and the science community, with access to funding and facilities for cost and feasibility studies.

3. The appointment of a scientist of outstanding achievement in the space sciences as Chief Scientist of NASA could be valuable; consideration should be given to the creation of a small science institute as part of the Chief Scientist's office.

4. There should be increased effort to obtain the services of academic scientists on leave to work at NASA headquarters.

8.6 SUMMARY

In this chapter, we have discussed manpower and funding in earth and planetary physics as a whole and have presented some details about various fields that contribute to work in this subfield: atmospheric sciences, geological sciences, ocean sciences, and space sciences. Discussions of priorities in NAS and government reports are cited and discussed.

It was beyond the scope and capability of this Panel to revise existing priority statements in the earth sciences. However, the breadth of our study offers an unusual view of a large area of a scientific activity, and we have noted and discussed a series of questions that resulted from this novel perspective and advocated the following courses of action:

In some areas of earth sciences, cost–benefit studies should be pursued with more energy and depth than has hitherto been the case.

We believe that the Geodynamics Project, when it has fully evolved, should receive enthusiastic support from relevant agencies, and we urge them to follow its development with a view to early implementation of its proposals.

We suggest to the Physics Survey Committee that it urge the appropriate agencies to review the following possible courses of action relating to the underemployment of new PhD's in physics:

1. To put immediate risk money into postdoctoral fellowships in universities, national centers, and federal laboratories specifically for transfers of young physicists into the earth sciences.
2. To consider procedures, possibly involving variations of Civil Service rules, to absorb research workers in the "holding pattern" created by postdoctoral fellowships in the earth sciences into permanent employment.
3. To accord high priority to those plans for expansion in the earth sciences that would profit from an intake of young physicists.

We recommend to the Physics Survey Committee that the National Academy of Sciences and the National Academy of Engineering institute a study of the rationale for a space science program in the broad context of U.S. science and technology.

We recommend to the Physics Survey Committee that the support proposed for space physics in the 1970 report on priorities for space research be regarded as minimal.

We recommend to the Physics Survey Committee that greater effort be devoted to problems at the interface between the responsibilities of NASA and the user agencies.

To bring about a closer relationship between NASA management and active research scientists, we recommend to the Physics Survey Committee the following measures:

1. *Greater reliance should be placed on standing disciplinary subcommittees in which experience can be gained over many years. We suggest that these subcommittees be used more frequently as an informal forum for the discussion of future policy.*
2. *There should be more science planning groups, composed of representatives of both NASA and the science community, with access to funding and facilities for cost and feasibility studies.*
3. *The appointment of a scientist of outstanding achievement in the space sciences as Chief Scientist of NASA could be valuable; consideration should be given to the creation of a small science institute as part of the Chief Scientist's office.*

4. There should be increased effort to obtain the services of academic scientists on leave to work at NASA headquarters.

Appendix A: National Facilities

The following list provides examples of a broad range of institutions and programs in which large or expensive facilities are operated in the national interest and are available to competent members of the scientific community. The financial sponsors—the National Science Foundation (NSF), National Aeronautics and Space Administration (NASA), National Oceanic and Atmospheric Administration (NOAA), and Advanced Research Projects Agency (ARPA)—and estimated annual operating budgets appear in parentheses. In addition to these national facilities, earth scientists have access to a variety of other facilities operated by states and private institutions.

1. *National Astronomy and Ionospheric Center* (NSF, $2.5 million). The NAIC operates an incoherent scatter radar with capabilities in upper atmospheric physics and chemistry. Cornell University operates this facility, which is shared by aeronomers, radio astronomers, and radar astronomers.

2. *Global Atmospheric Research Program* (U.S. portion is funded principally by the Department of Defense, NASA, NOAA, and NSF, $10 million in fiscal year 1970, $18 million in fiscal year 1971). The GARP is a joint program of the International Council of Scientific Unions and the World Meteorological Organization to establish a worldwide weather observation system as input to mathematical models of the atmosphere for the purpose of improving predictions. Many different projects are supported under GARP. The most recent is BOMEX, to be followed by a more ambitious experiment in tropical Atlantic regions.

3. *Jicamarca Radio Observatory* (NSF, NOAA, Peruvian Government, $2.5 million). The JRO operates an incoherent scatter radar at the magnetic equator with capabilities in upper atmospheric physics and chemistry. Built by NOAA but now owned by Peru, JRO is operated by the Instituto Geofisico del Peru, with limited scientific support from a group of eight U.S. staff associates.

4. *Joint Oceanographic Institutions for Deep Earth Sampling* (NSF, $8 million). The JOIDES collects, logs, and stores cores from ocean beds for

analysis by a group of university-based scientists. The coring operation is managed by Scripps Institution of Oceanography through a subcontract with Global Marine Corporation.

5. *Planetary Sciences Division of Kitt Peak National Observatory* (NSF, $0.7 million). The PSD operates a sounding rocket program and provides rockets, launch and control, and assistance as needed in developing the scientific package. The KPNO is operated by the Associated Universities for Research in Astronomy.

6. *Large Aperture Seismic Array* (ARPA, $5 million). The LASA is a 200-km-diameter array of 525 seismometers in Montana, with the data telemetered to the Seismic Array Analysis Center run by the International Business Machines Corporation and also analyzed by the Lincoln Laboratories of the Massachusetts Institute of Technology.

7. *Lunar Receiving Laboratory of the Manned Spacecraft Center* (NASA, $1.9 million). The LRL is devoted to the preliminary examination, handling and distribution, and continuing studies by a resident staff assisted by visiting scientists of samples of the moon returned by the Apollo missions.

8. *Lunar Science Institute* (NASA, $0.75 million). The LSI brings university scientists in contact with NASA's Manned Spacecraft Center and its Lunar Receiving Laboratory. The LSI is operated by Universities Space Research Association and is physically located adjacent to the Manned Spacecraft Center.

9. *National Center for Atmospheric Research* (NSF, $11 million). The NCAR sponsors a large fundamental research effort, principally in the physics and dynamics of the earth's lower atmosphere. It manages aircraft, balloon, and computer facilities and includes the High Altitude Observatory (solar physics). The University Corporation for Atmospheric Research operates NCAR.

10. *Office of Polar Programs* (NSF, $10 million). With the logistic support previously supplied by the U.S. Navy, observational work is conducted in the Antarctic in geology, oceanography, and atmospheric sciences and will be conducted in the Arctic, especially in meteorology and glaciology.

11. *Office of Space Science and Applications* (NASA, approximately $500 million). The OSSA operates rockets, earth orbiters, and planetary missions coordinating scientific packages that observe the atmosphere of the earth and other planets, the earth's surface and its resources, and the surfaces of the moon and nearby planets.

Both in regard to the size of its budget and the range of its activities, the OSSA differs greatly from the other facilities and programs listed here. Activities that could be classified as national programs or facilities are the Rocket Program, the Data Center at Goddard Space Flight Center, and various large and small space projects, all of which are accessible to U.S. and foreign scientists on a competitive basis. Large projects include the Mars Viking, Outer Planets Missions, Proto-Large Space Telescope, Apollo Telescope Mount, Skylab, and High Energy Astronomical Observatory. Smaller programs include the Small Astronomy Satellites, Pioneers, Planetary and Atmospheric Explorers, Nimbus, Earth Resources and Technology Satellites, and the like.

12. *World-Wide Standardized Seismic Network* (NOAA, NSF, $0.25 million). The WWSSN operates some 115 stations around the world that send their daily seismograms to NOAA for microfilming. These film strips can be ordered from NOAA, which acts as a data center.

13. *Naval Research Laboratory Upper Air Physics and Rocket Spectroscopy Sections* (Department of Defense, NSF, approximately $1 million). Support from the NSF enables the laboratory to entertain proposals for rocket research from university scientists.

14. *National Oceanographic Instrumentation Center* (NSF, Department of Defense, NOAA, $1.9 million). Operated by the Navy. The Center develops and designs oceanographic instruments and offers extensive facilities for testing and calibrating equipment.

Appendix B: Data Centers

The following data centers are located in the United States and can be used as facilities by all workers in the field. (Data centers other than those in the United States are not included in this Appendix.)

INTERNATIONAL DATA CENTERS

Data Center	Types of Data
Aeronomy and Space Physics	
World Data Center A Upper Atmosphere Geophysics National Oceanic and Atmospheric Administration Boulder, Colorado 80302	Solar and interplanetary phenomena, ionospheric phenomena, flare-associated events, aurora, cosmic rays, airglow
World Data Center A Geomagnetism	Geomagnetic and magnetic phenomena, digitized

Data Center	Types of Data
Environmental Data Service, NOAA Boulder, Colorado 80302	geomagnetic data
World Data Center A Rockets and Satellites Goddard Space Flight Center Code 601 Greenbelt, Maryland 20771	Rocket and satellite
Meteorology	
World Data Center A Meteorology National Weather Records Center Asheville, North Carolina 28801	Meteorology and nuclear radiation, central collection and publication of *Meteorological Rocket Sounding Data*
Oceanography	
World Data Center A Oceanography NOAA Rockville, Maryland 20852	Oceanographic observations
Seismology and Geology	
World Data Center A Seismology Environmental Data Service, NOAA Boulder, Colorado 80302	Tabulations of reduced data obtained from seismograms in station bulletins, microseismic data in bulletins and reports
World Data Center A Tsunamis NOAA P.O. Box 3887 Honolulu, Hawaii 96812	Seismograms and mareograms from selected stations in connection with tsunamis
World Data Center A Upper Mantle Project Archive Environmental Data Service, NOAA Boulder, Colorado 80302	Reports giving the results of scientific investigations of the upper mantle that include: recent movements of the earth's crust, paleomagnetism, volcanology, geochemistry, properties of rocks under high pressures and temperatures, geothermics, deep drilling
International Tsunami Information Center NOAA, P.O. Box 3887 Honolulu, Hawaii 96812 (contiguous to WDC-A Tsunamis)	Reports and data on tsunamis

U.S. DATA CENTERS

Data Center	Type of Data
Aeronomy and Space Physics	
National Space Science Data Center Goddard Space Flight Center Code 601 Greenbelt, Maryland 20771 (contiguous to WDC-A for Rockets and Satellites)	Reduced data obtained from space exploration projects
Aeronomy and Space Data Center, NOAA Boulder, Colorado 80302 (contiguous to WDC-A for Upper Atmosphere Geophysics)	Solar and interplanetary phenomena, ionospheric phenomena, flare-associated events, aurora, cosmic rays, airglow
USAF Environmental Technical Application Center Building 169 Navy Yard Annex Washington, D.C. 20390	Environmental data from earth's surface to near space
Meteorology	
National Weather Records Center Asheville, North Carolina 28801 (contiguous to WDC-A for Meteorology)	Meteorological data, original records
Oceanography	
National Oceanographic Data Center Building 160 Navy Yard Annex Washington, D.C. 20390 (contiguous to WDC-A for Oceanography)	Oceanographic station data, bathythermograph data, current and drift data, biological data, geological–geophysical data
Smithsonian Oceanographic Sorting Center Smithsonian Institution Washington, D.C. 20560	Oceanographic, biological, and ecological specimens
Seismology and Geophysics	
National Geophysical Data Center Seismological Branch Asheville, North Carolina 28801	Seismograms on microfilm from over 115 stations of the World-Wide Standardized Seismic Network
National Oceanographic Data Center and the Smithsonian Oceanographic Sorting Center (listed above under *Oceanography*)	Geological data in ocean areas

Appendix C: Balance of Activities in the National Science Foundation Programs

In response to the needs of the science community, the NSF has supported three national centers: Kitt Peak National Observatory (KPNO) and the National Radio Astronomy Observatory (NRAO) in astronomy, and the National Center for Atmospheric Research (NCAR) in the atmospheric sciences; and, recently, it has agreed to take responsibility for the Arecibo Observatory for astronomy and aeronomy, previously supported by DOD. These enterprises have been an outstanding part of the NSF program, and their success has led to continued growth and to recommendations for the establishment of new centers. Current proposals include a new aeronomy facility and a high-pressure facility for geophysics.

These national centers have expensive plants to maintain, tenure and nontenure staff, and continuing research programs, all of which involve a long-term commitment of support by the NSF. Since the NSF is generally their sole source of support, it has a direct responsibility that has no parallel in university research and other activities funded under its disciplinary programs. In these activities, the NSF acts as a partner, often providing most of the funds, but without direct responsibility for tenure or the long-term health of the programs.

During the past decade, no particular problems have arisen because of this dual responsibility within the same budget structure. Indeed, it can be argued persuasively that the success of the national centers has been partially responsible for maintaining or increasing expenditures in other parts of the NSF budget. However, we believe it is important to maintain a careful balance of activities, in view of the possibility of future retrenchment or additional pressure on NSF programs resulting from the withdrawal of research support in the earth sciences by other agencies.

When agencies have a dual responsibility to in-house and extramural research, there is a predictable response to budgetary restrictions: namely, to withdraw first from the extramural activities. This reaction is natural and represents sound management practice. Nevertheless, it probably is not the best response in terms of the overall national science capability. The effect of current withdrawal of support for the universities by DOD, AEC, and other agencies could be serious, and it would be critical if the NSF were not providing steady and increasing support. If budgetary constraints were to force the NSF to withdraw support suddenly, as other agencies have done, irreparable harm could result. Therefore, we believe that caution must be exercised in the expansion of existing programs or the creation of new programs for which the NSF has direct managerial responsibility. A balance must be sought that allows for the possibility of future fiscal stringencies.

We cannot assume that, faced with fiscal restrictions, the NSF could act differently from other agencies. A mild precursor is the expenditure ceiling imposed on the universities since fiscal year 1969. This example is difficult to analyze because it did not result in an overall reduction in university expenditure. However, in specific cases, it did result in retroactive cuts, without warning, of as much as 40 percent, leading to substantial decreases in research activity. Possibly, some relief could have been afforded the atmospheric sciences programs by decreasing the inhouse activity in NCAR. However, as illustrated in Figure IX.16, the operation and program equipment budget of NCAR rose during this period ($9.8 million in 1968, $10.4 million in 1969, and $11.0 million in 1970).

In the creation of new centers, the NSF acts mainly in response to the expressed wishes of the science community. The responsibility for a sound and balanced program of research support, therefore, rests in part with this community, particularly the NAS. We suggest that the Academy devote increased effort to the review of committee recommendations, with particular attention to the long-term implications of their proposals with respect to the balance of science activities supported by the NSF.

References

1. Aristotle, *Meteorologica* (translated by E. W. Webster) (Oxford U.P., New York, 1923).
2. C. B. Boyer, *The Rainbow; From Myth to Mathematics* (Yoseloff, New York, 1959).
3. L. F. Richardson, *Weather Prediction by Numerical Process* (Cambridge U.P., Cambridge, 1922; reprinted by Dover, New York, 1965).
4. J. Smagorinsky, "Problems and Promises of Deterministic Extended Range Forecasting," Bull. Am. Meteorol. Soc. *50*, 286–311 (1969).
5. F. G. Shuman and J. B. Hovermale, "An Operational Six-Layer Primitive Equation Model," J. Appl. Meteorol. *7*, 525–547 (1968).
6. G. P. Cressman, "Public Forecasting–Present and Future," in *A Century of Weather Progress* (American Meteorological Society, Boston, 1970), pp. 71–77.
7. Committee on Atmospheric Sciences, *Weather and Climate Modification: Prob-*

lems and Prospects, NAS-NRC Publ. 1350 (National Academy of Sciences–National Research Council, Washington, D.C., 1966).

8. Space Science Board, *Planetary Exploration 1968–1975* (National Academy of Sciences, Washington, D.C., 1968).
9. Committee on Atmospheric Sciences, *The Atmospheric Sciences and Man's Needs: Priorities for the Future* (National Academy of Sciences, Washington, D.C., 1970).
10. Commission on Marine Science, Engineering, and Resources, *Our Nation and the Sea: A Plan for National Action* (U.S. Government Printing Office, Washington, D.C., 1969).
11. Space Science Board, *Priorities for Space Research 1971–1980* (National Academy of Sciences, Washington, D.C., 1971).
12. *Annals of the IQSY* (MIT Press, Cambridge, Mass., 1968–1970) (7 volumes).
13. Johns Hopkins University Center for Research in Scientific Communication, *The Role of the National Meeting in Scientific and Technical Communication* (Baltimore, Md., 1970), Rep. No. 5, p. 251, and Rep. No. 6, p. 293.
14. Committee on Meteorology, *Research and Education in Meteorology: An Interim Report* (National Academy of Sciences–National Research Council, Washington, D.C., 1958).
15. Committee on Atmospheric Sciences, *The Atmospheric Sciences 1961–1971*, Publ. 946 (National Academy of Sciences–National Research Council, Washington, D.C., 1962).
16. U.S. National Committee for the Global Atmospheric Research Program, *Plan for U.S. Participation in the Global Atmospheric Research Program* (National Academy of Sciences, Washington, D.C., 1969).
17. W. J. Maunder, *The Value of Weather* (Methuen, London, 1970).
18. Committee on Solar-Terrestrial Research of the Geophysics Research Board, *Physics of the Earth in Space: The Role of Ground-Based Research* (National Academy of Sciences, Washington, D.C., 1969).
19. Ad Hoc Panel on a New Upper Atmosphere Observatory of the Geophysics Research Board, *Upper Atmosphere Observatory: Criteria and Capabilities* (National Academy of Sciences, Washington, D.C., 1971).
20. Geophysics Research Board and Division of Earth Sciences, *Solid-Earth Geophysics: Survey and Outlook*, NAS-NRC Publ. 1231 (National Academy of Sciences–National Research Council, Washington, D.C., 1964).
21. International Union of Geodesy and Geophysics–International Union of Geological Sciences *ad hoc* Committee on the Long-Range Program of Solid Earth Studies, *Final Report*, July 1969 (Reprinted in Trans. Am. Geophys. Union, Feb. 1970, pp. 152–159).
22. Committee on Earthquake Engineering Research, *Earthquake Engineering Research* (National Academy of Sciences, Washington, D.C., 1969).
23. Committee on Seismology, *Seismology: Responsibilities and Requirements of a Growing Science–Part II, Problems and Prospects; Part I, Summary and Recommendations* (National Academy of Sciences, Washington, D.C., 1969).
24. National Aeronautics and Space Administration, *The Terrestrial Environment: Solid-Earth and Ocean Physics*, NASA CR-1579 (U.S. Government Printing Office, Washington, D.C., 1970).
25. U.S. Geodynamics Committee, Geophysics Research Board, "Geodynamics Project: Development of a U.S. Program," Trans. Am. Geophys. Union *52*, 396–405 (May 1971).

26. Space Science Board, *Planetary Astronomy: An Appraisal of Ground-Based Opportunities*, Publ. 1688 (National Academy of Sciences–National Research Council, Washington, D.C., 1968).
27. Space Science Board, *Planetary Exploration 1968–1975* (National Academy of Sciences, Washington, D.C., 1968).
28. Space Science Board, *Lunar Exploration: Strategy for Research 1969–1975* (National Academy of Sciences, Washington, D.C., 1969).
29. Space Science Board, *The Outer Solar System* (National Academy of Sciences, Washington, D.C., 1969).
30. President's Science Advisory Committee, *The Next Decade in Space* (U.S. Government Printing Office, Washington, D.C., 1970).
31. NASA Astronomy Missions Board, *A Long-Range Program in Space Astronomy* (National Aeronautics and Space Administration, Washington, D.C., 1969).
32. Space Science Board, *Venus: Strategy for Exploration* (National Academy of Sciences, Washington, D.C., 1969).
33. Space Science Board, *Outer Planets Exploration, 1972–1975* (National Academy of Sciences, Washington, D.C., 1971).
34. Division of Engineering, National Research Council, *Useful Applications of Earth-Oriented Satellites* (National Academy of Sciences, Washington, D.C., 1969).
35. Agricultural Board, National Research Council, *Remote Sensing: With Special Reference to Agriculture and Forestry* (National Academy of Sciences, Washington, D.C., 1970).
36. "NASA Cocks Snooks at Academy Advice," Nature *230*, 142–143 (1971).

X

Physics in Chemistry

PANEL MEMBERS

R. STEPHEN BERRY, University of Chicago, *Chairman*
WILLIAM GODDARD, California Institute of Technology
RICHARD E. MERRIFIELD, E. I. du Pont de Nemours & Co., Inc.
JOHN ROSS, Massachusetts Institute of Technology
JEROME D. SWALEN, International Business Machines Corporation

LIAISON REPRESENTATIVES*

BERNARD G. ACHHAMMER, National Aeronautics and Space Administration
RICHARD J. KANDEL, Atomic Energy Commission
ROLF M. SINCLAIR, National Science Foundation
DONALD K. STEVENS, Atomic Energy Commission

* In general, no agency was represented on the Committee or any of its panels by more than one person at a given time.

Contents

X
PHYSICS IN CHEMISTRY

1 Physics, Chemistry, and Their Interface

This report deals with the interaction between physics and one of the disciplines most closely associated with it—chemistry. Interaction between these disciplines occurs in most of the subfields of each but is strongest in an interface that includes chemical physics and several other closely related subjects.

We shall first describe what we mean by physics, chemistry, and the interface areas and explain the terms we use in our later discussion. We then turn to an examination of the flow of information from one discipline to the other, the impact one has on the other, and some of the developments arising specifically from strong interactions between them. The next chapter is devoted to subjects outside the scope of either chemistry or physics that now derive necessary and important information from both subjects and especially from the interface. Subjects that are just beginning to use such information, or that can be expected to emerge soon and to require both physical and chemical concepts, are then considered. We discuss next the current manpower situation and projected needs. The arguments in this section provide a basis for a discussion of support and priorities. We conclude with a brief consideration of basic and applied science that develops some of the points presented in the report and reflects the views of many of the poeple who work at the interface.

It is difficult and probably unnecessary to formulate precise definitions to distinguish physics from chemistry. Those characteristics that differentiate the two disciplines relate more to goals and attitudes than to explicit content. Physicists strive to establish general laws common to all of nature. A great deal of the research that formerly was done only by physicists is now conducted also by chemical physicists, electrical engineers, and mechanical engineers. Chemists study phenomena to improve their understanding of the differences among substances. For example, the chemist differentiates substances through their physical properties, such as reflection of visible radiation; he determines their mass by their behavior in a gravitational field; and he studies in detail the processes that take place in chemical reactions.

Other distinctions between these disciplines have been suggested. Some people restrict the definition of chemistry to studies concerned with extranuclear matter, so nuclear-level structure would fall outside chemistry, but chemical shifts in Mössbauer spectra would be within chemistry. Much of what has been called solid-state physics falls into the intermediate area between physics and chemistry because of its concern with substances and its attention to the distinctions between these substances. We have assigned many subjects to the intermediate area on the basis of the viewpoints and attitudes associated with them and held in common by people from both chemistry and physics working in these subjects.

In attempting to locate the subjects at the interface, we were aware that physicists and chemists in academic institutions tend much more than those in other institutions to be overly concerned with defining their own traditional subjects. A major consequence has been a corresponding overconcern with specific training in narrow professional fields. Serious problems arise from this narrow professionalism, particularly for a scientist wanting to change fields or to broaden his interests in other fields. These problems result in part from the different vocabularies employed in various fields to describe essentially the same things. Another aspect of the problem is the existence of social and intellectual relationships that sometimes make it more difficult for work to be recognized or evaluated properly when a scientist works in a new field as a stranger than when he works in a field in which he is well known and well established.

To define the subject of strongest interaction between physics and chemistry, we use examples rather than an exhaustive listing of research topics. We attempt to establish only the rough outlines of the interface.

A substantial part of the interface often is called chemical physics. (It is interesting to note that in at least one industrial research laboratory, the name "molecular physics" seemed to be more socially acceptable, for purposes of hiring, than "chemical physics." Many young people working in the interface subjects prefer to be identified as chemical physicists rather than physical chemists. We shall comment further on this point in a later chapter.) Subjects usually considered part of chemical physics include theory pertaining to atomic and molecular wavefunctions; spectroscopy of atoms and molecules, from the radio-frequency range to some aspects of gamma-ray spectroscopy; chemical kinetics and collision processes, including hot-atom chemistry, radiation damage, and atomic and electronic collision processes at energies up to kilovolts or perhaps, tens of kilovolts; the entire field of the structure of liquids; polymeric molecules; statistical mechanics of both equilibrium and transport processes; molecular crystals, but only certain highly selected properties of other types of crystals; constitutive thermodynamic properties of all phases of matter except perhaps superfluids, superconductors, and plasmas; the use of lasers to study properties of matter (for example, nonlinearities); and the mechanism of laser action, excluding the optical properties of coherent light. In the United States, these topics are included within the scope of chemical physics; however, several of these subjects are considered pure physics in Europe and Japan.

Subjects that fall into the interface but are frequently or generally excluded from chemical physics include the following: properties of ionic, metallic, and covalent crystals and other periodic structures; amorphous materials such as glasses and semiconductors; the origin of the elements and several other aspects of nuclear physics, including many aspects of nuclear spectroscopy and nuclear reactions; and superconductivity, particularly with regard to its microscopic level of interpretation and the development of superconductive materials having specific properties.

Although chemistry and physics contribute to many subjects and subfields that are not always regarded as chemical physics, other subjects appear to suffer unduly slow development because of inadequate communication between these disciplines. The development of materials having specific desired properties—for example, crystalline and amorphous solids having particular optical or electrical properties, liquids with particular viscous properties, or gas mixtures with special capabilities for energy transfer—provides an illustration. As later examples will indicate, the communication between physicists and chemists dealing with properties of materials appears to be much more frequent and effective in industrial laboratories than it is in most universities.

2 The Influence of Physics and Chemistry on One Another

In some respects the relationship between physics and chemistry is highly asymmetrical. One illustration of this is the flow of manpower from subjects traditionally in one discipline to subjects associated with the other. A second illustration relates to the influence of one discipline on the other.

Although people trained as physicists in some instances practice what they call chemistry, movement from one discipline to the other generally is from chemistry, especially physical chemistry, into chemical physics or physics. Often these physical chemists and chemical physicists move into subjects in physics that pose challenging and important problems but that currently receive little attention from the physics community. Possibly, the tendency for persons trained in chemistry to work on problems formerly regarded as belonging to traditional physics reflects the changing nature of chemistry and the delay between the time an approach is first used to solve chemical problems and the time when it is well assimilated into the chemical literature. The trend also could be a reflection of a social "pecking order" among scientists.

In regard to the influence of the content of one discipline on the other, we are clearly in a period in which most of the basic theoretical framework for chemistry and most of the chemist's physical methods for experimental research trace their origins to physics. Quantum mechanics, statistical mechanics, and thermodynamics—the entire theory of the extranuclear structure of matter—are essentially physical laws that underlie any theoretical interpretation of chemistry. Many (but not all) chemists now feel that the theory of the chemical bond has been reasonably well explained in terms of quantum mechanics. Fundamental and productive insights continue to occur in chemistry as a result of the application of quantum mechanical methods to chemical problems. One particularly powerful tool, recently developed, is a set of interpretive and qualitative rules governing the spatial arrangements preferred by reacting species. These rules often permit the synthetic chemist to select from several possibilities a specific path by which he can control in detail the geometric arrangement of atoms within a molecule. He can, for example, find methods for synthesizing only the active form of a chemotherapeutic drug, instead of having the greater part of his product produced in inactive forms. Formerly, only natural biosynthetic processes exhibited this kind of efficiency.

Experimental chemistry has been deeply influenced by physical methods. Spectroscopy, from the radio-frequency region of nuclear magnetic resonance, through microwave, infrared, visible, ultraviolet, and x-ray, and even some aspects of gamma-ray spectroscopy, is vital to the chemist. Nuclear resonance has become an exceedingly powerful tool for analyzing chemical structures, because the response of most nuclei to this probe is highly characteristic of the immediate environment of that nucleus. For example, one can count the numbers of hydrogen atoms in each of many kinds of sites, even in rather complex molecules. The method is being developed and used now for the study of the structures of protein molecules. Electron spin resonance in the microwave region permits scientists to probe the reactive site (as well as the structure) in the large class of highly active intermediates known as free radicals, the species responsible for the formation of many polymers. Microwave and infrared spectroscopy have enabled the chemist to study the sizes and shapes of molecules by letting him see how they rotate and vibrate. They also make possible the recognition of characteristic reactive parts of molecules for purposes of identification and monitoring. Ultraviolet and x-ray spectroscopy are probes for the electrons within the molecule; they permit the chemist to study, for example, the chemical bonds and the colors of substances. Chemists have applied lasers in much of their work, including photochemistry, the study of energy conversion, and the probing of the energy levels of bound electrons. Other techniques from physics have equally broad use; mass spectrometry and x-ray crystallography, for instance, are now necessary tools for the chemist.

A characteristic pattern is associated with the assimilation of any physical method into chemistry. When the method is first discovered, a few chemists, usually physical chemists, become aware of chemical applications of the method, construct their own homemade devices, and demonstrate the utility of the new tool. At some point, commercial models of the device are put on the market. These are sometimes superior, sometimes inferior, to the homemade machines in terms of their ultimate capabilities to provide information. However, the commercial instruments generally are easier to use and far more reliable than the homemade devices. The impact of the commercial instruments is rapidly felt, is often very far-reaching, and sometimes virtually revolutionizes a field. Chemists with the new instruments need not be concerned with developing the principle of the device; they are free to devote their efforts to extracting the useful chemical information that application of the device affords. This pattern characterizes the development of optical, infrared, and radio-frequency spectroscopy, mass spectroscopy, and x-ray crystallography. Any reasonably well-trained chemistry student has learned enough of each of these to be able to recognize their capabilities and, perhaps with a bit of brushing up, to apply them to appropriate chemical problems.

Other experimental methods that have had great impact on chemistry include radioactive tracer and activation techniques, which are employed in both qualitative and quantitative analysis. These methods use the properties of the nucleus in a way analogous to that in which infrared and ultraviolet spectroscopy use the properties of electrons and atoms. Some of the most dramatic applications of these methods are in interfaces between chemistry and other disciplines. For example, radioactive dating methods have become an integral part of archaeology. Neutron activation and x-ray fluorescence also are used in archaeology as well as in the study of the fine arts and in criminology. The first analyses of the composition of the moon were accomplished by automatic analytical devices on the Surveyor spacecraft that determined the composition of the lunar surface by measuring the response to radioactive components of the devices.

When a physical method has proven its ability to give accurate and reliable chemical data, it usually finds a place as a monitoring probe, often in industrial process control. The oil industry could hardly exist as we know it now without real-time mass analysis of the contents of its reaction vessels.

The flow of new physical methods into chemistry continues. Electron spectroscopy—the analysis of electrons knocked from the molecules of a sample by light, other electrons, or other forms of excitation—is a new and powerful tool for studying the way electrons hold a molecule together. The first commercial instruments are just appearing, and a new kind of chemical information is rapidly being accumulated. Another current example is the use of low-energy electron diffraction for the study of surfaces. Slightly less developed but very promising is the use of very-long-wave infrared spectroscopy, actually interferometry, to study surface properties, such as epitaxial growth. Chemists have used this method less than have physicists studying properties of solid materials; it offers an example of the delay that can result from weak interaction between fields. Still another example is the application of Mössbauer spectroscopy to the study of dynamical properties of ions dissolved in water and other solvents.

Theoretical concepts, as well as hardware, flow from physics to chemistry. Modern methods for dealing with interactions among electrons—problems of the electronic structure of atoms and molecules and of scattering of electrons by atoms and molecules (as well as molecule-molecule interaction)—have strong ties to the many-body methods developed in the context of nuclear physics. In some cases, very closely related methods were developed independently by scientists, some of whom identified themselves as chemists, and others as physicists.

Chemical information, other than theory, feeds back into physics, often in the form of specific information about substances having some particular desired properties. The development of the ruby and neodymium lasers, dye lasers, and liquid lasers in general was made possible by the availability of the kind of information that chemists now collect routinely. The subject of properties of materials also includes problems of preparation and purity. The development of solids for microcircuitry required new chemical research into methods of purification and analysis. Properly functioning microcircuit elements of one large computer were developed only when the physical and chemical problems of ion migration under conditions of high current flow and methods to achieve an exceedingly high degree of purification had been solved.

An example of an active area of materials preparation, in which physics and chemistry interact strongly, is the new field of bubble memory devices. It is to be hoped that these devices, based on the behavior of orthoferrite materials, will lead to a new generation of computer memories with fewer failures and faster access time.

One type of interaction between chemistry and physics falls outside the categories we have mentioned. It relates to the compilation of data. The American Petroleum Institute has produced extensive tables of physical properties that, in effect, correlate properties of substances with their molecular structure. The National Bureau of Standards and the Joint Army, Navy, NASA, Air Force Interagency Propulsion Committee (JANNAF) prepare similar compilations of other kinds of data. The National Bureau of Standards also has a system to provide standard samples for those who need them. It is possible that other industrial groups would find it worthwhile to participate in and encourage data-collection enterprises.

3 Fields Deriving Information from the Physics–Chemistry Interface

In this chapter, we cite examples of fields outside physics or chemistry in which a combination of ideas from physics and chemistry either is now generating rapid development or can be expected to stimulate development in the next few years. In some cases, the contributions take the form of very specific information; in others, they are more in the nature of general viewpoints and techniques that are common in physics and chemistry but are just being recognized as useful in another field.

The knowledge obtained from basic studies in chemistry and physical chemistry often finds early application in engineering and technology. Some types of studies having particularly great potential for technological development are corrosion, the synthesis of new materials such as crystalline polymers, the nature of surfaces, heterogeneous catalysis (including the study of enzyme reactions), and thin films. All of these types of research require concepts and methods from both physics and chemistry. The techniques employed, such as low-energy electron diffraction, electron microscopy, electron microprobes, and electron spectroscopy, are being assimilated into physics and chemistry. An example will show the interaction of chemistry and physics and the close relationship to technological developments. A manufacturer recently developed a material for photocopying—an organic charge-transfer photoconductor. The development of this material can be traced from the basic concepts of the quantum theory of the chemical bond, through the chemical idea of a charge-transfer complex, to the recognition of photoconductivity. Current knowledge in basic physics and chemistry permitted the research for this material to advance rapidly to the point where it was apparent that an electron acceptor was desired, thus reducing the amount of expensive search time required.

The study and management of the oceans, the earth, and the atmosphere rely heavily on both physics and chemistry. In oceanography, for example, the turnover of water in the oceans can be understood only through a study of both hydrodynamic flow (traditionally a subject of physics) and molecular flow (as much a part of chemistry as of physics). The carbon dioxide balance in the air and the absorption of carbon dioxide by the oceans are critically important and difficult basic problems in physics and chemistry and in biology, meteorology, and environmental management, as well. The management of fisheries and water resource management in general require inputs from physics and chemistry in addition to the more obvious ones from biology. Air resource management is perhaps one of the clearest examples of an activity in which chemistry and physics play an integral role. Whether one is concerned with control of dangerously high pollution levels or with long-range planning and legislation, air quality models are essential to the analysis required in any program of air resource management. Such models depend on knowledge of the behavior of masses of gas (fluid dynamics), the reactions within the gas (aerochemistry), and the influence of the earth's shape, motion, and position on the way this gas behaves (meteorology). Only within the past two years have the first models begun to appear that incorporate all these components at a level of accuracy high enough to be useful for an air resource management program. One of the major drawbacks has been our limited knowledge of the basic physics and, especially, the basic chemistry of polluted air.

Modeling is an approach that is natural to a physicist or chemist as a result, for example, of his familiarity with the equations of transport. The general treatment of linear systems or, more important, nonlinear systems and their instabilities is characteristic of both disciplines. Such approaches are useful in the development of social and economic models and probably will be extremely helpful in predictive analyses of secondary consequences of potential new technologies.

In regard to social or economic modeling, the physicist or chemist can make a special type of contribution. The constraints incorporated in a model are of two kinds, although the models now in use do not make this distinction. One type of constraint is essential and unavoidable, depending only on the laws of thermodynamics and other equally general laws. The other type of constraint is basically a limitation resulting from the current state of technology. By separating the first type from the second, one should sometimes be able to determine the ultimate limiting behavior of a system—the condition it could approach (but never reach) if its technology were optimal. Such an analysis could be made, for example, to determine the optimal long-term choices for supplying sufficient water to an industrial nation. A similar kind of analysis, using population biology instead of thermodynamics, could be used to determine the amount and pattern of insecticide to apply to maintain the long-term productivity of an agricultural area and the health of the animal species in the associated food chain.

Other fields in which physics and chemistry are making contributions are the structure of biomolecules, including the analysis of the action of enzymes, the process of photosynthesis, the structure of cell walls, and the action of nerves; new methods of materials preparation, such as ion implantation, discharge methods for petroleum cracking, and surface curing by electron bombardment; geophysics, including the origins of volcanism, the chemical origin of rocks and minerals, and, of course, the composition of the planets; and the upper atmosphere and astrophysics, including the entire complex set of reactions of our own atmosphere as well as the problem of the formation of molecules in interstellar space.

4 The Population and Support of the Physics–Chemistry Interface and Related Fields

4.1 THE POPULATION AT THE INTERFACE

Having reviewed some of the areas in which physics and chemistry interact and stimulate new work, we examine now the distribution of research personnel in these fields and the approximate turnover in manpower. The figures we use are not precise, largely because the definition of the interface is not precise. The figures include estimates of the upper and lower limits on the number of people working in subjects at the interface. We have tried to make all the estimates of these limits on self-consistent bases, so that the upper limits, for example, are based on the total picture of the interface, including chemical physics and other parts of overlap. The lower limits are based only on the number working in chemical physics in academic, governmental, and industrial institutions.

The total interface contains between 4000 and 8000 scientists at the PhD level. (The larger figure includes solid-state physics.) This estimate is based on figures from the National Register of Scientific and Technical Personnel and on an independent estimate derived from the number of subscribers to the *Journal of Chemical Physics.* In chemical physics alone, the number of scientists is some 3000 to 4000. Of these, about 250 hold academic positions in chemical physics and about 500 hold academic positions in some other subfield included in the interface area.

The 1968 National Register recorded 8466 scientists working in physical chemistry, 4836 of whom had PhD's. The Register category "physical chemistry" is composed of an aggregation of detailed specialties. A scientist is classified as working in a field on the basis of his own statement of a detailed specialty that most nearly corresponds to the subject matter with which he is principally concerned in his work. Table X.1 presents the breakdown of the Register's physical chemistry aggregate in terms of the specialties and of the major scientific affiliation—American Chemical Society (ACS) or American Institute of Physics (AIP)—of the respondents. Only 391 physicists were working in physical chemistry compared with a total of 8466 working in this field. Many chemical physicists, of course, are doing work in subfields outside the scope of the Register definition of physical chemistry and would be included in, for example, the solid-state or fluid-physics subfields.

Of the AIP respondents to the National Register, 591 PhD's and 417 non-PhD's identified themselves as chemical physicists. Comparing these numbers with the people working in physical chemistry, we find that of some 1000 chemical physicists, 391 are working in the specialties listed in Table X.1 and the rest are distributed among other subfields.

The field of highest degree provides another estimate of the number of AIP physicists with some involvement with chemistry. There were 543 PhD's in chemistry represented in the AIP Register population, 295 of whom received their degrees in physical chemistry. (These respondents are not necessarily the same people who identified themselves as chemical physicists in other parts of the Register questionnaire.)

According to the National Research Council Office of Scientific Personnel, about 500 PhD's are produced each year in "Chemistry, Physical" (459 in 1968 and 554 in 1970). These physical chemists are produced by chemistry departments; no record is kept of "Physics, Chemical" as a subcategory of physics department PhD production. For purposes of compari-

TABLE X.1 Chemists and Physicists Working in the Various Specialties Comprising Physical Chemistry in the 1968 National Register of Scientific and Technical Personnel

Specialty	Members of ACS	Members of AIP	Total
Catalysis and surface chemistry	784	15	800
Chemical and phase equilibria	142	8	150
Chemical kinetics, gas phase, and photochemistry	461	26	487
Chemical kinetics, liquid phase	270	6	276
Colloid chemistry	351	7	358
Crystallography	287	32	320
Electrochemistry	877	25	902
Energy transfer and relaxation processes	103	16	119
Flames and explosives	153	24	177
Fused salts	87	0	87
High-temperature chemistry	225	5	230
Ion exchange and membrane phenomena	142	3	145
Isotope effects	31	0	31
Liquid state and solutions; electrolytes and nonelectrolytes	182	13	195
Molecular spectroscopy	598	45	643
Molecular structure	152	14	166
Nuclear and radiochemistry	307	3	310
Polymers in bulk; morphology, phase transitions, rheology, and mechanical properties	678	67	745
Polymers in solution; thermodynamics, hydrodynamics, and spectroscopy	273	12	285
Quantum and valence theory	290	14	304
Radiation and hot-atom chemistry	241	11	252
Solid-state chemistry	374	18	392
Thermochemistry, thermodynamics	407	3	410
Statistical mechanics	74	0	74
Other	583	24	608
TOTAL	8072	391	8466

son, it is worth noting that, in 1970, physics departments awarded 400 solid-state doctorates, 470 elementary-particle and nuclear physics doctorates, and 730 doctorates in 11 other physics areas, including 160 in "Physics, Other." The total number of physics PhD degrees reported in 1970 was 1600.

The migration of physicists into chemistry, or any other discipline, has never been very large. On the National Register, only 118 of 29,000 doctorate chemists obtained their PhD's in physics. Adding all physicists with degrees in chemistry (543) to the physics PhD's now identified with chemistry (118) gives a total of 661.

Larger numbers for the population at the interface are derived from American Physical Society activities. The Division of Chemical Physics has 1800 members. There is no breakdown of PhD's. The entire American

Physical Society is represented on the National Register by 16,300 of its membership of 28,000. In this portion, 24 percent were nondoctorates; there is no other clue to the composition of the membership in terms of highest academic degree level.

Journal subscriptions are difficult to interpret. The *Journal of Chemical Physics* has 2232 domestic subscribers who belong to at least one of the AIP member organizations. (The total number of subscriptions is 6300, but this figure includes libraries and foreign subscriptions.)

The throughput—the replacement rate required to keep existing positions filled—is estimated by the Panel to be currently about 6 ± 1 percent per year in the interface. This figure is based on several estimates, such as the normal throughput in a typical industrial laboratory. The figure appears to be fairly firm, except in years such as 1970, when the rate of hiring is held at a lower level. Consequently, we estimate that the total number of replacement personnel required at the PhD level is 240–280 per year for the overall interface and about 180–240 per year for the more strictly defined field of chemical physics.

If we assume that the level of scientific research and development support during the past 5 years represents, approximately, an upper limit of the fraction of its gross national product that this nation is prepared to spend for such activities, and if we also assume that the actual funding level will keep pace with inflation and that cutbacks in research and development spending by the Department of Defense will be shifted to support of other subject areas, we reach the disturbing conclusion that we are producing PhD's to work in the chemistry–physics interface at a rate significantly higher than that at which we are able to use them. Whether our figures are precise is not important; what is important is that the rate of production is much higher than the requirements or capabilities for assimilation now or in the next few years.

4.2 RESPONSES TO THE PROBLEM

Responses to the problem of overproduction of manpower range from shifts in the methods of academic training to shifts in the patterns of financial support. First, it is clear that a new PhD will have to be more flexible in his choice of position than were his predecessors in the early 1960's. The doctorate will have to represent a level of ability in identifying, setting up, and solving difficult challenges rather than the passage through a craftsman's apprenticeship, as it often has been in recent years. This suggests that future training at the predoctoral level should place increased emphasis on introducing the student to individual responsibility for defining problems, with less emphasis on participating in team research in which problems are already well defined. Problems demanding team research should be dealt with in most instances by semipermanent and postdoctoral scientific staffs rather than predoctoral students.

Once through his doctoral work, the young scientist should be prepared to learn the specific background necessary for his new job. The formal course curricula in university physics departments could aid the student in making this transition by being increasingly flexible in regard to the inclusion of courses in molecular physics, fluid dynamics, thermodynamics, and fields peripheral to physics, such as chemistry, geology, and astronomy. In essence, we are suggesting that a doctoral training program should instill an attitude different from the one that is now prevalent. It is our impression that many young scientists have been reluctant in recent years to consider work outside the specialty of their doctoral training. (We have also heard criticisms that employers often look for PhD's with specialized training to fill a highly specific slot.) The doctoral student should not expect to continue a career in the specialty of his doctoral research; rather, he should expect to move from his graduate work into any of a wide variety of subareas in physics or tangential to it. To foster the growth of such an attitude will require realization and acceptance by faculty members that urgent practical problems can provide challenging questions for basic research.

4.3 SUPPORT OF WORK AT THE INTERFACE

We turn now to the problem of future financial support for the physics–chemistry interface. Our examination of this problem is based on the number of people now working in the interface, the recent rate of growth of the field, and the estimates of annual cost for support of one scientist in an academic or industrial laboratory.

The total population of the interface is between 4000 and 8000 scientists at the PhD level, when the interface is defined broadly enough to include those parts of solid-state physics, nuclear studies, cryogenics, and other fields in which the chemistry–physics interaction is strong. We estimated that the more narrowly defined field of chemical physics contains about 3000 to 4000 PhD scientists.

The current annual cost of support for one of these scientists is approximately \$60,000. Hence, the current annual spending for the entire interface would be between \$200 million and \$400 million, and the total support for chemical physics, between \$150 million and \$200 million. Of these totals, the amount allocated to academic institutions merits particular attention, because it probably comprises the overwhelming part of the non-mission-oriented funds, which take the form of federal research grants. There are about 500 PhD's in academic positions in the entire interface, so that the total support of academic work at this interface is on the order of $25 million. About half as many people can be called chemical physicists; thus the annual support for academic work in this field is currently about $12.5 million.

Population trends in the interface suggest that maintenance of a constant population requires the replacement of abut 6 percent of the PhD's in the field each year. However, the rate at which PhD's have been generated in recent years has been roughly twice that required to maintain a steady-state population. Whether this growth rate of 6 percent, in addition to the 6 percent replacement rate, can be expected to continue during the next few years is uncertain. However, if the scientists now being trained—that is to say, those new graduates who have just entered the field and those students who are still in graduate school—are to be supported, a 6 percent increase in funds would be needed for each of the next 4 years. This figure does not include inflation, which would require an additional 5 percent per year. Therefore, maintenance of full employment of the people now working in the field or being trained to enter it would require an annual increase in support of about 11 percent for each of the next few years. Such a level of support would correspond to maintenance of the status quo in terms of present population and current amount of support for each scientist.

We have assumed that further increases in funding (apart from increases to meet the costs of inflation) are unrealistic. We must base our future planning, for at least the next few years, on the level of support that might be expected rather than on a hypothetical ideal number of people who should be working in the field. This estimate includes only a 5 percent annual increase in funds to offset inflation.

To maintain the status quo under these conditions will require shifting research efforts into some of the subfields mentioned in Chapter 3 and decreasing the number of doctoral students in the field to about half the present number. We have been concerned primarily with how best to accomplish this goal while maintaining the present vigor of our training programs. The interface is attracting a highly creative group of young people and is in an extremely productive stage of development. We must continue to attract creative young scientists to the field to maintain the desired replacement rates in universities. Furthermore, the store of basic research information must be constantly increased if new applied science and technology are to thrive.

The maintenance of quality must be the primary concern during the anticipated period of approximately level, or decreasing, support. The recent rapid expansion and creation of doctoral training institutions has resulted in an overproduction of about a factor of two. Peripheral considerations, such as local geographical interests, should not override the necessity for maintenance of high quality as we adjust to this overcapacity. We suggest that universities carefully assess their doctoral programs. Less prestigious graduate schools could find themselves with few, if any, doctoral students. Employers presumably will select new PhD's from the schools with stronger reputations and programs of high quality; therefore, some schools probably will find it very difficult to place their students. In view of these considerations, it would be advisable for universities, especially those with new or weak graduate programs in the physical sciences, to examine these programs critically to determine whether to continue them. There could be strong reasons for some schools to adopt different programs, such as in-service training in applied fields or courses oriented toward teaching, possibly without the requirement of a doctoral research thesis.

We strongly urge that, before any general new funding policy is adopted or recommended, a critical evaluation be conducted of the "centers of excellence" program to estimate the utility of institutional funding, the role of such institutions with regard to both established and new scientists, and the capabilities for growth in the physical sciences in U.S. universities. We recognize that this program is one important factor that contributed to our present overproduction of manpower. Experience with it emphasizes the necessity of assessing the potential drawbacks as well as the advantages of future programs.

This Panel believes that federal support for students should be offered predominantly through fellowships to individuals. In addition, in view of the increased breadth in training and point of view that we advocate, we propose that these fellowships be awarded on the basis of broad general subject areas rather than narrow and highly specific ones. The support of a minor fraction of students, and of most postdoctoral workers, could be retained in the research grant and contract programs. If universitywide or departmentwide support is instituted, we urge that such grants be subjected to regular external peer-group review.

Should the dollar level of support remain approximately constant, this circumstance would imply a contraction of the size of the effort at the interface, either in the number of people or in the productive capability of each man, or in both. Our firm recommendation to the Physics Survey Committee, in this case, is to support excellence at an adequate level rather than to try to maintain the current broad base of the field at a reduced level. Funds spread too thin may well be wasted. It is more productive to determine the amount of money available and the amount necessary to support one man adequately and then to support as many as possible of those doing work of the highest quality.

5 Views on Basic and Applied Science

We regard the relationship between basic and applied science as an organic system in which all students are trained at a basic level, then enter either basic or applied work, as they choose, with the greater number of opportunities in applied areas. High priority should be given to fostering recognition of applied science as a valid intellectual pursuit for many active scientists. To promote this recognition, we suggest that scientific meetings include sessions devoted to applied areas (as American Physical Society meetings do now), that review articles in applied areas appear periodically in appropriate journals, and that seminar series in university physics and chemistry departments include lectures in applied areas. In addition, some universities should examine the possibilities and drawbacks of various types of joint university–industry training programs.

At the same time that we urge the breakdown of prejudices and barriers between pure and applied research, we must keep in mind a fundamental principle in setting priorities in research and development. Basic science, at least in the subject covered in this report, generally is more efficient, in terms of information per dollar or man-hour spent, than is applied science. Applied science, in turn, is far more efficient in these respects than is development. We cannot afford to reduce spending at the basic level, for the basic level of research is the appropriate place for continued exploration of novel, high-risk ideas. Support of basic research is the natural way to sustain the flow of new concepts and approaches.

Progress, especially at the basic level, is erratic, unpredictable, and dependent entirely on the talents of the individuals in a field. Consequently, support—particularly support of basic research—should be directed far more often toward people than toward projects. In regard to development, the reverse is true; here, we can set national or institutional goals and direct support toward those goals.

Finally, each scientist must develop a sensitivity to the significance and implications of his work. That is to say, each scientist should consider the future implications of his work as well as its immediate cogency. With such an attitude, we may avoid the trend toward medieval scholasticism that now threatens us.

XI
Physics in Biology

PANEL MEMBERS

ROBERT SHULMAN, Bell Telephone Laboratories, Inc., *Chairman*
VICTOR P. BOND, Brookhaven National Laboratory
FRANKLIN HUTCHINSON, Yale University
SEYMOUR H. KOENIG, International Business Machines Corporation
ROBERT LANGRIDGE, Princeton University
ALEXANDER MAURO, Rockefeller University

LIAISON REPRESENTATIVES

EUGENE L. HESS, National Science Foundation
ROBERT W. WOOD, Atomic Energy Commission
RICHARD S. YOUNG, National Aeronautics and Space Administration

Contents

XI
PHYSICS IN BIOLOGY

1 The Nature of the Field

Although most of the panels are concerned with subfields of physics, we will describe the interaction of physics with biology. To describe this interaction, we must examine the flow of manpower, ideas, and procedures through the interface as well as the subject matter of the interface. The title, Physics in Biology, implies a flow from left to right. It is convenient to divide the subject into three parts: the flow of physicists into biology; the flow of physics into biology, which includes the flow of ideas, techniques, and equipment, often through the intermediaries of engineering and chemistry; and the interface called biophysics in which the powers of physics are merging with the problems of biology in new ways that require nonstandard combinations of skills from both disciplines.

The goal of biophysics is to be biologically useful. When it is successful, it merges with other fields of biology, such as biochemistry and molecular biology, a major goal of which is to understand, in molecular terms, how genetic information is transmitted. The following example illustrates the contributions of biophysics to molecular biology. In 1944, it was shown that DNA, rather than proteins, contained the genetic information with which molecular biology is principally concerned. In 1953, J. Watson and F. Crick, working in the Cavendish Laboratory, developed from the accumulated chemical studies of DNA and a consideration of its biological function an interpretation of M. Wilkin's x-ray studies in terms of the now famous double helix. Since then, molecular biology has developed rapidly, while biological function has become so interwoven with the structure of DNA that the term "structure and function" has become a platitude. At the same time, and in the same laboratory in Cambridge, the crystal structures of the proteins were first being determined by x-ray crystallography, *with a parallel influence on enzymology and biochemistry. In all these* cases, the isolation of the molecules and the definition of their biological functions were accomplished after about a century of chemical research. The physicists who determined the crystal structures were attacking a biologically important problem. Their boldest and most original step was their starting assumption that these large biological molecules had unique structures that could be determined by x-ray crystallography. Physicists are accustomed to this kind of simplicity in science; assuming such simplicity is a reflection of their previous training and research style. The revolutionary nature of their findings depended on the originality of their assumptions. However, given their background as physicists, trained in x-ray crystallography under W. L. Bragg and interested in biology, the direction their efforts would take was almost predetermined. Thus, in these illustrative and illustrious cases, physicists and physics created the exciting subfield of biophysics, which has now been merged with biochemistry and molecular biology.

In the immediate future, the determination of the structure of large biological molecules will continue to be an active field of research. In addition to x-ray crystallography, other physical techniques, such as nuclear magnetic resonance, electron spin resonance, Mössbauer studies, and optical studies, are being used more frequently. These other techniques, when used for structural studies, are often applied to complement x-ray data by giving information on a finer scale or in solution. This research is directed toward structural determinations of larger molecule aggregates, that is to say, membranes and membrane-mediated enzyme systems; ribosomes, which are the site of protein synthesis, composed of nucleic acids and protein, with a molecular weight of $\sim 10^6$; mitochondria, the membrane-bound volume in which the chemical energy of nutrients is converted to more usable forms by electron-transfer reactions; and the photosynthetic unit in which photons are converted to chemical energy. In all of these systems, scientists are trying to understand biochemical functions in terms of the structure of the molecules and the physical interactions among them. Beyond the structures of isolated biological molecules lie the complicated questions of intermolecular interactions, which should challenge physical methods for many years.

When we examine structure at finer levels than the molecular, it is quite clear that quantum mechanical understanding of the electronic structure of certain parts of biological molecules will become increasingly important.

The advances made through electronic understanding of the molecules of interest to chemistry and solid-state physics show the promise of this approach. Recently, as the experimental molecular physicists have studied biological molecules with the goal of understanding their electronic properties, the amount of systematic data has approached the volume needed for theoretical synthesis and advances. This synthesis could lead in the future to a larger role for theoretical studies. Theorists' contributions to molecular biophysics have heretofore been very small because, unlike the best experimenters, they generally have not learned enough biology to enable them to ask good questions. In their approach to biology, computational techniques or solutions in closed form have often served as the basis for questions. One way in which theoretical physicists have contributed to modern biology is through the development of models, as described in Appendix A.

In contrast, physicists who have immersed themselves in molecular biology have contributed major simplifying ideas based on fundamental physics and the research style of physics. Examples are S. Benzer's topological study of a gene, Crick's postulate of a genetic code followed by the "wobble" concept, which allowed freedom to the intermediates, and M. Delbruck's insistence on orderly phage genetics. (These examples are discussed in Appendix B by means of the individual biographies.)

2 Institutions in the Subfield

Academic research in this subfield, in contrast to research at institutes, carries the burden of departmental responsibilities. Two means have been used to broaden the scope of academic research. In many cases, temporary departments of biophysics were established and made into formal departments when it seemed desirable. This practice has worked very well in some universities and less well in others. In the most successful cases, the biophysics departments so established became departments of modern biology and often lost much of their initial chemical or physical orientation. However, the more physics-oriented aspects of biophysics still flourish in university departments of this sort, generally under a title such as "Biophysics and Molecular Biology." Advanced electron microscopy research, new techniques of radioactive tracers, magnetic resonance studies, and x-ray crystallography have prospered in these departments and have profited from their interactions with the more biology-oriented activities.

In practice, these departments do not interact much with physics departments, although physics courses are increasingly important to their students. In this respect, it is generally felt that physics departments do less than chemistry departments to accommodate such students, among whom are an increasing number of premedical students who desire some physics courses. Revision of the physics courses to accommodate the needs of these students would be one way in which the interactions of physics and the life sciences could be strengthened.

This kind of increased flexibility on the part of physics departments and increased interaction with biology also would strengthen collaborative research among different disciplines. Collaborative interdepartmental research is the second way in which academic departmental limits have been overcome. However, support for biophysical research from within the department is very important. Consider, for example, the case of a tenured professor of physics sufficiently interested in biology and free of commitments to devote time to learning a new field. Even in this favorable case, the amount of support offered by the department in terms of new appointments, student requirements, and general acceptance of this type of work can affect the development of the work decisively. Younger men usually will not risk their chances for advancement by switching fields of interest, which is unfortunate, because it is easier for them to learn a new field. Tenured appointments for younger men are particularly sensitive indicators of the real interest of a physics department in biology; such appointments have been very scarce. Among the reasons for this scarcity are the difficulty of evaluating their work in a physics department and questions about the appropriateness of such work in a physics department. When biochemistry has been conducted in physics departments, it has looked particularly polished, possibly because of the buffeting it has received. In several cases, this kind of informal interaction has profited considerably from the housing of physics, biology, and chemistry departments in the same or adjoining buildings.

Problems pertaining primarily to instrumentation probably could be studied in a physics department with a minimum of interaction. Although some techniques, such as electron microscopy, hold the promise of future applications in biology, in the past, few significant contributions to biology were made so directly. The biological applications must be very obvious before a physicist, who is somewhat removed from the discipline, has sufficient incentive to build a better instrument for a particular biological application. There are many cases in which physical instrumentation has been revised or extended for biological applications; however, this work generally has been done by instrumentation companies, especially in recent years. In some cases, such as spectroscopy, diffraction, and microscopy, the commercial instruments now used in biology were developed for the larger chemical physics markets. In other cases, however, such as centrifugation and particle counters, instruments were originally developed specifically for the biological market.*

Interdisciplinary research institutes represent another popular approach to broadening the scope of academic research. Such institutes have been established with participants from established academic departments joining those of the institute. Interdisciplinary institutes require rather unusual conditions that approach the more unified multidisciplinary environment sometimes found in governmental, industrial, and national laboratories. In these, the departmental barriers seem less constrictive, and it has been easier to organize groups in which people with different skills work on different aspects of a biological problem.

In the smaller liberal arts colleges, the introduction of biophysics or biophysical chemistry research as joint ventures of the physics, chemistry, and biology departments is particularly appealing for several reasons. There are many more worthwhile areas of research in biophysics and biophysical chemistry that can be undertaken on a very limited budget than there are in physics. There also are more worthwhile opportunities for undergraduate student participation in such research projects. Finally, because of the participation of three small departments, it is easier to generate a research effort of the required critical mass.

3 Interaction of Nuclear Physics with Biology and Medicine

Rather than presenting a general description of the widespread influence of physics on biology and medicine, we will discuss in some detail the influence of one subfield, nuclear physics, since radiation and health physics accounts for about half of the physicists listed by the American Institute of Physics as working in biological fields.

Biological and medical uses of nuclear physics have followed closely on developments in this subfield, and virtually every sector of biology and medicine has been strongly influenced, if not revolutionized, by its im-

* Recent instrumental advances made by physicists in the particularly important field of electron microscopy are discussed in Appendix C.

pact. Only selected examples are included in this discussion, with emphasis on those applications having most pertinence to medicine. Four areas are treated briefly: diagnosis, therapy, research, and radiation effects.

3.1 DIAGNOSIS

Radioactive isotopes have contributed enormously to improving diagnosis in general, and a large number of radioisotopes are now in routine use. Isotopes commonly used include ^{131}I, ^{125}I, ^{59}Fe, ^{113m}In, ^{99m}Tc, ^{51}Cr, ^{57}Co, ^{60}Co, ^{75}Se, ^{85}Sr, ^{197}Hg, ^{32}P, and ^{198}Au. These are used for visualization of the thyroid, brain, liver, lung, kidney, pancreas, spleen, heart, bone, and placenta and for a variety of physiological tests in which the rate of disappearance or rate of uptake of a particular labeled substance reflects the function of a given organ system. In 1966, there were some two million administrations of labeled compounds to patients, and the rate of use has grown rapidly. These isotopes are employed routinely in practically every hospital in the United States.

An example of the many uses for a specific isotope, technetium-99*m*, will illustrate its value as well as the physical characteristics of isotopes that make them useful. Technetium is taken up selectively by the thyroid. The energy of the gamma rays from technetium, approximately 0.14 MeV, are optimal for external detection and thus for imaging purposes. The half-life of the isotope, approximately 6 hours, is long enough for ready use; on the other hand, it is short enough that a minimal radiation dose is delivered to the patient in the course of obtaining the necessary diagnostic information. The isotope is a pure gamma emitter, and the ratio of information gained per radiation dose delivered is maximal. In other applications, the element can be used to label proteins, colloids, and other biological compounds, making the isotope useful for a wide variety of diagnostic procedures, including scanning of the brain, bone marrow, kidney, liver, and lungs. Thus, technetium-99*m* is an "all-purpose" isotope, which, because of its desirable physical characteristics, is now in extremely wide routine use.

Increasing attention is being given to the application of very short-lived isotopes, particularly ^{11}C, which should be very useful because of the enormous potential for incorporation into a wide variety of biological compounds, with consequent extension of the range of diagnostic procedures. The very short half-life, some 20 min, markedly limits the time for synthesis of the isotope into the desired compound and the time available for use. Thus the source of the isotope (an accelerator), facilities for rapid synthesis, and clinical facilities must be in close proximity. A team of physicists, chemists, and clinicians is required for effective application. Use of this isotope, although in its relatively early stages, is growing rapidly and has great promise.

Another rapidly developing diagnostic procedure involves the determination of the entire amount of a given element, for example, calcium, in the body by means of activation analysis. The entire body is exposed to a beam of fast neutrons, which thermalize in tissue and are captured by the element. The patient is then placed in a whole-body counter, and the total amount of the given element is determined. The entire procedure can be accomplished with the delivery of only a small fraction of a rad to the patient.

Critical to the development of diagnostic procedures involving radioactive isotopes is instrumentation. Although some procedures can be carried out with rather crude detectors, scanning procedures demand increasingly sophisticated equipment to obtain greater resolution in time and space. A wide variety of scanners have been developed that not only increase the resolution but permit acquisition of enormous amounts of information in a short time with minimal amounts of the isotope (thus keeping the radiation dose to a minimum). The use of computers is mandatory to reduce the data and provide an effective display in a short time.

3.2 THERAPY

Isotopes are now used widely in radiotherapy in a variety of ways. High-intensity external sources, such as cobalt-60 and cesium-137, have in many instances replaced x rays for routine radiotherapy. The depth–dose characteristics of radiations from these sources are more favorable than those of most x-ray beams, and the units have the advantage of ease of operation and maintenance. In addition, beta emitters such as strontium-90 are used externally for treatment of some superficial lesions.

Radioisotopes also are used internally for radiotherapy. The classic example is radium, applied in the form of needles, for implantation directly into tumor tissue. A variety of isotopes, including ^{60}Co, ^{192}Ir, and ^{125}I in the form of needles or wires are now used for this purpose. In addition to cost considerations, the radiations given off and the half-lives are more suitable for a variety of purposes than those of radium.

Physiological localization of radioactive isotopes is used in some forms of therapy—for example, ^{123}I and ^{131}I for treatment of hyperthymidism and thyroid tumors and ^{32}P for treatment of some diseases of the bone marrow. These procedures represent optimal therapy only in a relatively few situations.

Accelerators have contributed very significantly to the improvement of radiotherapy. Early accelerators allowed the very significant transition from the use of relatively low-energy x rays for radiotherapy to the use of "supervoltage" x rays, permitting the delivery of a relatively large dose to the tumor in depth, with minimal dose to the intervening normal tissues. Electron accelerators such as the betatron have permitted additional distinct improvement in the therapeutic ratio, that is, the dose-to-tumor/dose-to-normal-tissue ratio.

Currently, there is much interest in the use of accelerators to produce beams of fast neutrons for radiotherapy. The rationale is that all tumors quickly develop small foci of poorly oxygenated, or hypoxic, cells. These hypoxic cells are markedly resistant to damage by x-ray or gamma radiation but are much more susceptible to damage by neutrons or other densely ionizing radiations. Although a variety of reactions and neutron spectra might be used, the approximately 14-MeV neutrons from the deuterium–tritium reaction are optimal in terms of penetrating characteristics and density of ionization. The procedure is experimental, and several years will be required to evaluate its efficacy. Isotopic sources of neutrons are also being used, with the same rationale. Transplutonic elements such as californium emit fast neutrons spontaneously, and ^{252}Cf is suitable in terms of half-life. The material can be fashioned into needles or wires and used for implantation therapy in much the same fashion as radium sources. This approach is also experimental, and its ultimate contribution remains to be determined.

Other accelerator-produced radiations have promise in radiotherapy. One of considerable interest is a beam of pure π^- mesons. These particles behave essentially as do electrons, in terms of dE/dx losses, throughout most of their range in tissue. At the end of their range, as a result of nuclear interaction, a variety of densely ionizing particles, including protons, alpha particles, fast neutrons, and stripped atomic nuclei, are released. The resulting depth–dose curve permits the deposition of a relatively high dose of densely ionizing radiation essentially selectively to the tumor in depth, with the delivery of a relatively low dose of sparsely ionizing radiation to intervening normal tissues. At present, beams of π^- mesons are inadequate in terms of intensity and purity for radiotherapy trials. Machines are becoming available in Vancouver (TRIUMF) and in Zurich that may be adequate for initial therapy trials. The output of the π-meson factory* under construction at Los Alamos will be adequate for this purpose.

Also potentially useful in radiotherapy are beams of very-high-energy stripped nuclei of Z of the order of 6 or above. These beams can be

* Los Alamos Meson Physics Facility (LAMPF).

shaped to obtain an improved therapeutic ratio, and the densely ionizing character of the beam renders hypoxic cells more susceptible to the radiation.

Instrumentation has also played a key role in the development of therapeutic procedures. The improved ability to deposit energy selectively in the region of the tumor has increased the demand to better determine the confines of the tumor and to tailor the overall application to maximize the therapeutic ratio. Thus, sophisticated treatment planning has become a necessary part of therapy, and the use of computers is both intensive and extensive.

3.3 RESEARCH

The ability, provided by isotopes, to label and trace biological compounds without physiological or pharmacological effects from the traces has opened areas of investigation that formerly were inaccessible. Essentially all research in biology and medicine now involves the use of either radioactive or stable isotopes.

Accelerators have been used to some degree in biological and medical research. Radiation provides a clean method to perturb parts of an organism selectively. For example, beams of protons have been used to produce laminar lesions in the brain for the purpose of determining the function of the destroyed area. Portions of the bone marrow have been selectively perturbed by radiation to study hemostatic mechanisms and factors that determine the rate of regeneration of damaged tissues. Cardiac pacemakers and artificial hearts powered by nuclear sources could well become practical in the next several years.

3.4 RADIATION EFFECTS

Extensive knowledge of the effects of radiation is necessary for radiation therapy, radiation protection, military and space medicine, and the treatment of accidental overexposure. Increasingly, radiation sources are becoming a part of modern living, and it is necessary to know in some detail the degree of effect as a function of dose and dose rate. For extended missions in space, we must know if radiation exposure might seriously injure the astronauts or actually compromise the success of the mission. Thus, the effects of radiation in space, including the effects of very-high energy heavy galactic particles, must be studied.

An enormous amount of effort has been devoted to the study of radiation effects; radiation can be characterized as the most extensively studied and best understood environmental toxic agent to which man is exposed. However, a great deal remains to be learned. It is important to attempt to understand the action of radiations at a molecular level, since only through such understanding can we adequately evaluate radiation effects under the myriad conditions of exposure.

4 Training in the Subfield

In the numerous ways described in the previous chapters, we have seen how necessary the contributions of physics are to modern biology. Furthermore, there is every reason to believe that these contributions will continue to increase rapidly, on a relative if not absolute scale, depending on the level of support. These activities are particularly rewarding in a personal sense, first, because of the exhilarating intellectual challenge of this diverse and fast-moving field and, second, because of the humanistic aspects of biological research, with its long-range goal of improving the human condition. Physicists have moved into biology and made significant contributions. Many more would like to do so. The problem is how to facilitate this transition.

As background for this discussion, we must mention several characteristics of the flow of physicists into biology. First, it is numerically much smaller than we had expected. Second, and corollary, the physicists who do change fields stand out rather prominently. Third, although physicists who work at the interface can retain their identity as physicists, those who do are not likely to make any serious contributions to biology.

We will separate physicists interested in biology into two groups: those who adopt biological goals, in E. Pollard's terms "surrender" to biology, and those who retain their physical goals but include the subject matter of biology within these goals. These two groups have many things in common; therefore, in discussing their training requirements, we will consider them both, distinguishing between them only when necessary.

A study of changes of subfield of specialization among 4000 physicists between the years 1960 and 1966 showed that only 13 of these physicists moved into biology in this period. Since nearly half of the total number of U.S. physicists were followed in this survey, we can conclude that about four physicists per year, who are sufficiently qualified in the discipline to be included in the Physics Section of the National Register of Scientific and Technical Personnel, change to biology. On the other hand, about 50 physics participants in each of three National Register surveys—1962, 1964, and 1966—were biophysicists. Although this number was approximately constant for each of the surveys, the population varied, with about a 40 percent turnover every 2 years. Nearly all of those who left biophysics returned to other subfields of physics. Of the 13 physicists who moved into biology in this period, 8 came from condensed matter, 3 from nuclear physics, and 1 from elementary-particle physics, with the remaining man not accounted for. These data deal with well-established physicists; therefore, they indicate midcareer transitions. It is the Panel's feeling that an increasing number of students go into biology after taking an undergraduate degree in physics. We have no data yet to support this view. However, we see that biophysics, considered as a means for physicists to enter biology, has a very high reflection coefficient. From our personal experiences, we believe that a large fraction of those studying or working in biophysics were open to the possibility of making a larger commitment to biologically relevant research. In general, they found a way of using their background in physics so that for a time, without learning much biology beyond the requirements of their immediate problem, they were able to function as biophysicists. When the initial problem on which they were working was solved, most had not ventured beyond the barriers that separated them from the rest of biology. With their mission finished, they returned, somewhat discouraged, to a more traditional branch of physics in which they knew that they could contribute meaningfully. Both the data obtained in the study and our own experience convince us that a barrier exists—not to entering biophysics but to moving from biophysics into biology. There are several ways to train the individual so that, if he wishes, he can surmount this barrier. The first, of course, is to educate him more broadly as a biophysicist so that there are more opportunities available—more biological pathways beckoning.

It seems to be well established that physics, chemistry, and mathematics are either learned well at an early stage in one's education or not learned at all. That is to say, the transition from physicist to biophysicist is very common, the transition from biologist or medical man to biophysicist is far less usual. Thus, the best foundation for a career in biophysics should be a broad base in the areas of physics, chemistry, and mathematics. The growing tendency for intense specialization in physics departments has made a bachelor's degree in physics less attractive as a base for a career in biophysics, and the increasing physical sophistication in chemistry departments has enhanced the attractiveness of a bachelor's degree in physical chemistry.

Most new biophysicists now receive their PhD's in biophysics from a large university department in which the formal training tends to be allocated in the following way: one third to rectifying deficiencies in the undergraduate preparation of the student (for example, elementary biology for the physicists and physical chemists or organic chemistry for the physicists), one third to advanced courses in molecular biology, and one third to specialist courses such as biochemistry and x-ray crystallography.

No graduate-level physics courses are normally taken. The complexities of biological systems are such that it is now frequently impossible to treat them physically from a fundamental point of view or to make detailed and physically sophisticated calculations or experiments.

A second well-established route into biophysics is to get a PhD in physics, chemical physics, or physical chemistry and then move into biology. This procedure would be the obvious route for one who is interested, for example, in the quantum mechanical properties of large molecules. This path requires several years of postdoctoral research, including opportunities for self-expression, and is thus equivalent to a second PhD. This transition is accomplished much more easily by chemists than by physicists, because modern molecular biology and biochemistry are contiguous with modern chemistry. However, in the cases in which physicists have made the transition, the scientific results often have been more distinctive and original than are those resulting from the less difficult transitions from chemistry.

At its most extreme, the originality of a physicist's approach completely changes a field of biology. Past examples are easy to find: Pasteur, after receiving his degree, under Biot, in the optical rotatory power of molecules, followed up his interesting observation that organisms were able to separate racemic mixtures into optical active parts. Other examples appear elsewhere in this report.

Where might we expect the next revolutionary changes in biology to take place as a result of the discoveries or techniques of physics? They could relate to an even more microscopic quantum-mechanical view of life processes or to computer techniques that will rival and illuminate biological activities or to an understanding of the neural processes, including the brain (see Appendix D) or to prosthetic devices. Our imagination can run wild, and, on the basis of past accomplishments, we can be sure that in some way our hopes will be fulfilled.

For the present, we must find ways to smooth the individual physicist's path. First, courses in biophysics specifically designed for the physicist and intended to bridge the gaps between these two fields could be taught in physics departments. Second, a slight increase in the degree of freedom accorded students in a physics department is desirable. But this depends on and is less important than the third point, which is that until physics departments desirous of encouraging biophysics are willing to promote young people to tenure positions on the basis of their work in biophysics, biophysics will not flourish as a subfield of physics. Without this step, there will be no sizable increase in biophysical research activity, and training arrangements can be made almost on an individual basis.

To enter biophysics, a midcareer training opportunity, such as a fellowship or sabbatical leave, should be supplemented by an additional one-year fellowship. Two years are barely sufficient to achieve professional competence in a new field. Fortunately, these midcareer fellowships can be supplemented by intensive summer school training such as that offered at Woods Hole or Cold Spring Harbor in different aspects of modern biology, including serious laboratory work. However, these courses and others will be sufficient only if accompanied by a change in research activities, best accomplished by a change in environment. Finally, meetings or lecture series, arranged to present the views of biology to physicists, or vice versa, must be judged, as they usually are, in terms of their social or cultural worth to the participant or in terms of their ambience, for, in terms of the re-education required, they are too superficial.

As previously mentioned, physicists working in biology tend either to adopt the goals of biology or to retain the goals of physics. The discussion in Chapter 3 of the applications of physics and biophysics to health and medicine in the particular field of nuclear medicine exemplifies the latter pattern of participation. Most of the research in this field can be performed by well-trained physicists who, in collaborative efforts and while furthering a biological goal, retain their identities as physicists. On the other hand, we also have discussed the few individuals who move from physics more completely into biology, very often in a highly creative way. These patterns of participation are, of course, extreme examples on a continuum, and creativity will not be found exclusively at either end. However, we do not believe that creativity is uniformly distributed. To stimulate maximum creativity, educational programs geared to the type of biophysicist one is hoping to produce should be developed.

5 Manpower, Cost, and Facilities

The number of people identified with physics in biology or biophysics depends on how the interface is defined. The number of physics participants in the 1970 National Register of Scientific and Technical Personnel who indicated physics in biology as the subfield in which they were working was 465, or 1.3 percent of the 36,336 physicists who took part in the survey. Sixty percent (277) of the 465 held PhD's, representing 1.7 percent of all physics PhD's. The median age of the physics-in-biology doctorate-holders in 1970 was 37.2 years, approximately the same as that for all physics PhD's—37.4 years.

If the population of the interface is based on membership in the Biophysical Society, then it increases to about 2500. The Society was founded in 1957, and its growth has been extremely rapid.

Data on employing institutions and principal work activities of the 465 physicists working in biology who participated in the National Register survey are available. These data indicate that nearly half the PhD group (46.3 percent) and about a third (35.0 percent) of the non-PhD's worked in colleges and universities. The "other" category in the Register questionnaire included hospitals and medical schools and was the one indicated by approximately one fourth of the physics-in-biology PhD and non-PhD groups (24.4 percent and 27.4 percent, respectively). A fairly high percentage of the non-PhD's (19.7 percent) reported industrial employment. Table XI.1 presents these data.

Basic research and teaching were the principal work activities reported by the physics in biology PhD group, as Table XI.2 shows. The non-PhD's were less involved in basic research and more involved in applied research than were the doctorates. The non-PhD's also reported substantial teaching and management responsibilities.

A comparison of work activity patterns for the physics-in-biology group and for the overall physics population indicates substantially heavier in-

TABLE XI.1 Employing Institutions of Physicists Working in Biology and of All Physicists

Employing Institution	Physics in Biology PhD's, $N = 274$ (%)	All Physics PhD's, $N = 16,248$ (%)	Physics in Biology Non-PhD's, $N = 157$ (%)	All Physics Non-PhD's, $N = 17,679$ (%)	Physics in Biology Total, $N = 431$ (%)	Physics Total, $N = 33,927$ (%)
College and university	46.3	50.6	35.0	30.0	42.2	39.9
Industry	10.9	23.4	19.7	30.2	14.2	27.0
Government	9.5	9.0	14.0	14.4	11.1	11.8
Research center	8.8	11.8	3.8	5.0	7.0	8.2
Other (including hospital and medical school)	24.4	5.2	27.4	20.4	25.5	13.1

TABLE XI.2 Principal Work Activity of Physicists Working in Biology and of All Physicists

Principal Work Activity	Physics in Biology PhD's, $N = 197$ (%)	All Physics PhD's, $N = 16{,}017$ (%)	Physics in Biology Non-PhD's, $N = 127$ (%)	All Physics Non-PhD's, $N = 17{,}001$ (%)	Physics in Biology Total, $N = 324$ (%)	Physics Total, $N = 33{,}018$ (%)
Basic research						
Primary	53.8	34.7	33.1	20.7	45.7	27.5
Secondary	22.8	24.9	5.5	7.2	16.0	15.8
Applied research						
Primary	9.6	16.8	18.9	20.8	13.3	18.9
Secondary	13.7	18.1	15.0	19.3	14.2	18.7
Design and development						
Primary	0.5	2.3	3.1	9.3	1.5	5.9
Secondary	1.5	6.6	11.8	14.9	5.6	10.8
Management						
Primary	7.6	16.0	14.1	16.0	10.2	16.0
Secondary	11.2	11.4	13.4	10.4	12.0	10.8
Teaching						
Primary	24.4	25.7	18.1	23.1	21.9	24.4
Secondary	26.9	18.2	19.7	7.6	24.1	12.8
Other						
Primary	1.5	2.4	10.2	8.0	4.9	5.3
Secondary	11.7	8.8	14.2	17.3	12.7	13.1
No response						
Primary	2.5	2.1	2.4	2.1	2.5	2.1
Secondary	12.2	12.1	20.5	23.3	15.4	17.9

volvement in basic research among those at the interface. In addition, about 10 percent more of those working at the interface indicated teaching as a principal work activity than did the overall physics population (46.0 percent compared with 37.2 percent).

The numbers of positions and graduates seem to be reasonably well balanced in the interface between physics and biology. There have been fewer positions in the past few years, as in all of physics, but the situation seems no worse in physics in biology than in physics in general.

As mentioned above, a training course in biophysics is particularly suitable for an emerging institution and would be better accepted there than in some of the older, well-established physics departments. Biophysics is similar to solid-state physics in regard to the optimum size of a research activity. With the exception of activities like instrumentation, which normally draws its support from much of a department, a biophysical activity would become more efficient with increasing size, up to the point at which three or four professors were involved. Equipment of a biochemical sort is most efficiently used in biophysical research when it is shared. For example, a preparative ultracentrifuge can be used by a biochemist to prepare his materials; it also provides a very popular means of following biological interactions. For the biophysicist, on the other hand, although this centrifuge is often useful for sample preparation, it is usually not used beyond that point, as the physical parameters are studied in other ways.

Another factor favoring larger research groups is the increasing cost of certain equipment used in biophysical research. For example, consider the cost of commercially available high-resolution nuclear magnetic resonance (NMR) spectrometers used to measure the proton NMR spectra of complicated organic molecules. In 1957, this technique was first used to study the spectrum of a protein, ribonuclease, with a molecular weight of ~13,000. During the past decade, stimulated in part by the possibility of extensive biological applications, this equipment has increased rapidly in price, as shown in Figure XI.1. At first it followed a semilog growth curve, but recently, because of the use of superconducting solenoids and Fourier transform computer techniques, the costs have been increasing more rapidly. At the same time that the equipment is becoming too expensive for an individual scientist to afford, its usefulness in biochemical research has been increasing rapidly because of improvements in sensitivity, resolution, and experience. Clearly, large institutional research activities, similar to the national laboratories, will soon be more useful in biophysical research.

Appendix A: The Role of Physical Models in Biology

The approach to biological problems taken by physicists is often different from that taken by people not trained in physics. One of these differences is the physicist's desire for a simple, comprehensive model, capable of providing a first-order explanation of a wide variety of observations. A physicist often is baffled by the biologist's insistence on the complexity of nature and the uniqueness of each result. It is, of course, equally unappealing to a biologist, struggling with the complexities of DNA replication, to be condescendingly informed by a physicist that the Ising model or enough molecular quantum mechanics would solve his problem. However, in the middle ground between these two extremes of oversimplification and non-generality lies the very fruitful, delicately balanced area of physical mod-

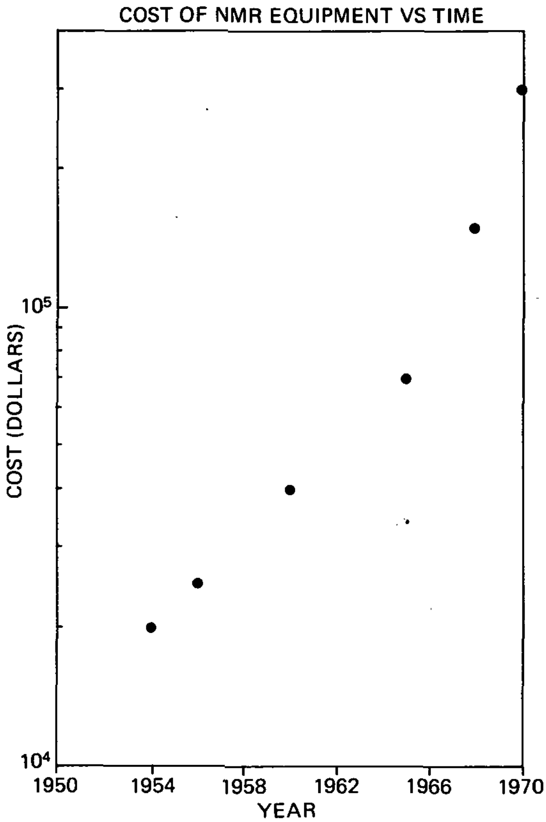


FIGURE XI.1 Cost of NMR equipment versus time.

els, based as always on experimental observations. For example, the concept of a genetic code proposed by physicists and theoretical chemists such as Crick, Orgel, Gamow, and Griffith was something that appealed to a physically trained mind. Various theoretical models were examined. One question that was proposed and answered was how much information had to be stored. The answer was that there were 20 amino acids that had to be coded by the DNA. Since DNA has only four possible bases to act as coding units, a minimum of three bases had to be required. Was the code overlapping? This was answered negatively by considering the known information of mutations that had been observed.

Another useful model is that of Monod, Wyman, and Changeux regarding allosteric proteins. The function of allosteric proteins with respect to one small molecule can be affected by other small molecules. They proposed a generalized molecular basis for feedback in biological molecules.

This is another example of a physical model that has helped biology enormously by stimulating many experiments and much thought about just what the crucial facts are.

In other directions, models have been proposed for biological structures on the basis of partial structural information. Among the most successful of these have been the Watson-Crick structure of DNA and Holley's cloverleaf structure of t-RNA in which, from the primary sequence, several hydrogen-bonded regions were proposed, the existence of which is now established. On a less detailed scale, the Danielli-Davson model of cellular membranes, with two outer protein layers sandwiching a double layer of lipids, has provided a rallying point for membrane structural studies for more than 30 years.

Appendix B: Physicists in Biology

One of the most remarkable features of the striking developments in molecular biology during recent years is the contribution made by individuals trained not as biologists but as physicists. Below are short biographies of four of the most prominent.

S. BENZER Born October 15, 1921, New York City, New York; PhD (physics), Purdue University, 1947; Professor of Biology, California Institute of Technology

F. H. C. CRICK Born June 8, 1916, Northampton, England; PhD (physics), Cambridge, 1953; Member, Staff, Medical Research Council, Laboratory of Molecular Biology, Cambridge

M. DELBRUCK Born September 4, 1906, Berlin, Germany; PhD (physics), Göttingen, 1930; Professor of Biology, California Institute of Technology

M. H. F. WILKINS Born December 15, 1916, Pongaroa, New Zealand; PhD (physics), Cambridge, 1940; Professor of Biophysics and Director of Biophysics Research Unit, Kings College, University of London

This list is not arbitrary—most biologists would include these names—but it could easily be lengthened.

The first question is to ask why these men left physics for biology: The reasons are diverse, but in searching for a common thread in their careers, Erwin Schrödinger's little book, *What Is Life?*[1] recurs. For three of these men, it was an important factor in their decision to go into biology.

CRICK Also he had read Erwin Schrödinger's book, *What Is Life?* This noted physicist had asked "How can the events in space and time which take place within the spatial boundary of a living organism be accounted for by Physics and Chemistry?" and had concluded that ". . . the obvious inability of present day Physics and Chemistry to account for such events is no reason at all for doubting that they can be accounted for by those sciences. . . ." What impressed Crick about the book at the time was that ". . . fundamental biological problems could be thought about in precise terms using the concepts of Physics and Chemistry." One was left with the impression that "great things were just around the corner."[2]

WILKINS I was . . . very much interested when I read Schrödinger's book, *What Is Life?* and was struck by the concept of a highly complex molecular structure which controlled living processes. Research on such matters seemed more ambitious than solid state physics.[3]

BENZER Delbruck first entered my life in the form of the chapter heading "Delbruck's Model" in Schrödinger's book, *What Is Life?* I read that book at an impressionable age while still a graduate student in pretransistor solid state physics at Purdue University.[4]

The last quotation gives a clue not only to Benzer's reasons for entering biology but also to Delbruck's relationship to Schrödinger's book. Rather than Schrödinger influencing Delbruck, Delbruck influenced Schrödinger, and, as Benzer remarks, one whole chapter by Schrödinger is devoted to Delbruck's model of the gene.

As for Delbruck himself:

My interest in biology was first aroused in Copenhagen by Bohr in connection with his speculations that the complementarity argument of quantum mechanics may have wide applications to other fields of scientific endeavor and especially in regard to the relations between Physics and Biology. The move to Berlin in 1932 was largely determined by the hope that the proximity of the various Kaiser Wilhelm Institutes to each other would facilitate the beginning of an acquaintance with the problems of Biology. This good intention eventually materialized. A small group of physicists and biologists began to meet privately (mostly in my mother's house) beginning at about 1932 . . . out of these meetings grew a series of papers (Wohl and Gaffron on photosynthesis and Timofeeff, Zimmer and myself on mutagenesis). A popularization of our paper of 1935 in Schrödinger's little book, *What Is Life?* had a curiously strong influence on the development of Molecular Biology in the late 1940's.[5]

Thus, although *What Is Life?* was of little direct scientific importance in that it did not solve any major scientific problems, it had an enormous indirect influence not only on some of the most eminent men in the field but also on many others. To follow the arguments of one of the greatest physicists and deepest thinkers of the time wrestling with the problems of *What Is Life?*—even to realize that he felt that the problem was ripe for a physicist's approach—had a deep and lasting impact on many physicists.

There were, of course, other motivations:

CRICK A more powerful stimulus than *What Is Life?* leading Crick to Biology was the "religious" one of an atheist who wanted "to try to show that areas apparently too mysterious to be explained by Physics and Chemistry, could in fact be so explained."[6]

WILKINS During the war I took part in making the atomic bomb. When the war was ending, I, like many others, cast around for a new field of research. Partly on account of the bomb, I had lost some interest in physics.[3]

Of the four, two (Benzer and Delbruck) abandoned completely what physics departments usually recognize as physics. Crick was initially, and Wilkins has remained, closer to the techniques of physics, but it was only because of the determined support of their laboratory directors (Bragg and Randall, respectively) that this kind of work ("dull" physics but exciting biology) was done in physics departments. (For an account of a contemporary physicist's reaction to Bragg's insight, see the article by Freeman Dyson.[7]) We wonder how many university physics departments would, even today, encourage the type of work done by Crick and Watson and by Wilkins. Depressingly few, one suspects. Perhaps this "hardening of the categories" afflicts all university departments. It is hardly their most laudable feature.

After his early work on DNA and protein structure, Crick strayed farther and farther from physics, and his important later contributions have been, like those of Delbruck and Benzer, to molecular genetics.

Although not pursuing what is commonly recognized as physics, these men brought with them an attitude of mind and a way of doing research that seemed more general in physics than in biology—the search for a simple system and, of course, a desire to seek explanations in terms of physics and chemistry. (Many biologists also felt the same way—the Beadle-Tatum one-gene–one-enzyme hypothesis is a classic example.)

Wilkins, more than the others in this group, remained close to physics in his techniques, beginning with optical microscopy (to which he made several important contributions) and culminating in his application of x-ray diffraction to DNA structure. X-ray diffraction, though a very powerful tool, is now a rather mundane branch of physics, but Wilkins, not a trained x-ray crystallographer, had the insight to choose DNA as worthy of attack by these methods, when other crystallographers, probably better trained than he, ignored it. (The sole exception was W. T. Astbury—another former physicist—who first took x-ray diffraction pictures of DNA in the late 1930's. But Astbury took x-ray photographs of almost everything biological that he could lay his hands on, and his *direct* contribution to the DNA work was small. However, his influence on the field as a whole was great, both in structural classification and understanding and in encouraging others, such as Wilkins, to believe that important and useful information could be gained from studying even these poorly ordered bits, pieces, and extracts of animals and plants.)

These outstanding examples of physicist-turned-biologist may have little relevance for the average physicist who makes such a move, but it is likely that the motivations are similar. Perhaps the striking feature of the most successful is their willingness to abandon the narrow view of physics and to tackle the new problems, not as problems in applied physics but as problems in science.

A quotation from Benzer provides a fitting conclusion for this section: "The best way to have fun in Science is to do something you are not trained for."[8]

REFERENCES TO APPENDIX B

1. E. Schrödinger, *What Is Life?* (Cambridge U.P., Cambridge, England, 1944; reprinted 1967).
2. R. Olby, "Francis Crick, DNA and the Central Dogma," Daedalus *99*, 942 (1970).
3. M. H. F. Wilkins, "Molecular Configuration of Nucleic Acids (Nobel Prize Lecture)," Science *140*, 941 (1963).
4. S. Benzer, "Adventures in the rII Region," in *Phage and the Origins of Molecular Biology*, J. Cairns, G. S. Stent, and J. D. Watson, eds. (Cold Spring Harbor Laboratory of Molecular Biology, 1966), p. 157.
5. M. Delbruck, private communication (biographical notes written for the Royal Society).
6. R. Olby, *ibid*, p. 943.
7. F. Dyson, "The Future of Physics," Physics Today *23*, 23 (1970).
8. S. Benzer, private communication (letter dated July 17, 1970).

Appendix C: Electron Microscopy

The speculations of de Broglie on the wave nature of matter soon caused individuals to investigate the possibility of forming images with a beam of particles having a wavelength much shorter than that of light. Within a decade, the first crude electron images had been obtained, using a basic design that was essentially the electron analogue of the conventional optical microscope. Impressive engineering advances have occurred in the 40 years since the development of these first microscopes; now most users can employ the instrument as they do an automobile, with little concern for the physical science or engineering involved. Resolution, contrast, and sample thickness are interrelated in a complex way, but, in practice, resolution is effectively about 15 Å for very thin biological samples and less favorable for thicker ones.

About 10 years ago, a fundamental departure in design was made to exploit more fully the characteristics of the electron beam used to form the image. In this microscope, a focused beam of electrons scans the surface of an object. The image is formed by collecting the secondary electrons ejected from the surface of the object and using the signal to modulate a television raster synchronized with the scanning beam. The principal advantages of this arrangement are a considerable increase in depth of focus and an apparent illumination of the object in a way equivalent to that by which we view ordinary three-dimensional objects. The result is a striking three-dimensional quality in pictures taken at a magnification of 10,000 to 20,000.

The next major advance occurred only very recently: A. V. Crewe devised a scanning electron microscope of the transmission type that gives information not only about the number of electrons passing through a given element of the image (as with a photographic plate) but also about the energy lost by the electrons scattered from the object and about the direction of the scattering. These additional data have made it possible to extract more information about the object and to effect considerable improvement in resolution and contrast. It is now possible to obtain images of a single heavy-metal atom such as uranium with 5 Å resolution, and it is likely that single atoms of medium weight will be resolved within the foreseeable future.

Appendix D: Physical Basis of Nerve Activity

A striking example of the success of physical reasoning in elucidating a particular property of a biological cell is the analysis of the electrical state underlying excitability in the squid giant axon. Its virtually unique diameter (500–100 μm) enabled Hodgkin and Huxley in 1949 to conduct a series of fundamental electrical measurements that, in turn, made it possible to establish for the first time an adequate quantitative description of the electrical state associated with the nerve impulse. Both the design and execution of the experiments required a thorough knowledge of electronic circuits, in which feedback plays a crucial role, and of the theory of ionic electric currents. Moreover, the interpretation of the data demanded an ingenious mathematical analysis. This achievement was in large measure a product of Hodgkin's and Huxley's training in physics; it would scarcely have been possible otherwise.

One of the basic results of their analysis was that the ionic currents in the axonal membrane display strikingly nonlinear behavior, in that the conductances are voltage-dependent and time-variant. The property of nonlinearity immediately implies that in a single neuron and in chains of neurons—that is, in neural circuits—the elaborate calculations required to treat the excitable state generally can be performed only by means of modern computers. The statistical physics underlying the conductance changes in the cell membrane and at the junction between two cells—the synapse—is a problem requiring analysis of considerable sophistication.

Briefly, then, the subject of excitability has attracted physicists in the past and undoubtedly will continue to do so. It is true that progress in neurobiology will demand new advances in the biochemistry and ultrastructure of the neuron, but, in any event, the elucidation of physical mechanisms, as we have seen in the analysis of excitability in the squid giant axon, will continue to play a crucial role. While DNA, RNA, and other aspects of molecular biology have become familiar to many physicists, the subject of excitability, ironically, in view of the inextricable relationship of its language and content with those of physical concepts, is still relatively unknown to them. No doubt, as more departments of physics turn to topics in biophysics, the phenomenon of excitability will come to be appreciated as a rich and rewarding subject by an increasing number of physicists.

XII
Instrumentation

PANEL MEMBERS

NATHAN COHN, Leeds & Northrup Company, *Chairman*
ERNEST AMBLER, National Bureau of Standards
W. A. ANDERSON, Varian Associates
J. G. ATWOOD, The Perkin-Elmer Corporation
T. G. BERLINCOURT, Colorado State University
GEORGE DOWNING, Merck & Company
E. C. DUNLOP, E. I. du Pont de Nemours & Co., Inc.
ERWIN EICHEN, Ford Motor Company
J. W. FRAZER, Lawrence Berkeley Laboratory
R. S. GORDON, formerly with The Monsanto Company
F. W. LUERSSEN, Inland Steel Company
EMERY ROGERS, Hewlett-Packard Company
D. B. SINCLAIR, General Radio Company
W. E. VANNAH, The Foxboro Company

LIAISON REPRESENTATIVES*

MARCEL BARDON, National Science Foundation
ROBERT L. BUTENHOFF, Atomic Energy Commission
PAUL F. DONOVAN, National Science Foundation
JAMES H. TRAINOR, National Aeronautics and Space Administration

* In general, no agency was represented on the Committee or any of its panels by more than one person at a given time.

Preface

The creation of an Instrumentation Panel was suggested by Survey Committee member Raymond Bowers. A major purpose of the Panel was to identify dependency, where applicable, of contemporary commercial instrumentation development and application in the United States on basic research in physics, done here or elsewhere in the world.

It was originally intended that the Panel be relatively small in size, but subsequent consideration prompted expansion of the Panel, permitting broad and varied representation from commercial instrumentation manufacturers and users. Invitations to serve on the Panel were extended and accepted with the understanding that the views developed and reported would be based on information already in the hands of Panel members and their associates and that, in view of time limitations, no extensive research to develop additional ideas could or would be undertaken.

The contents of this report, which it is hoped will be helpful in the Physics Survey, reflect the consensus of Panel members and do not necessarily represent views held by their respective organizations.

The following organizations and individuals assisted members of the Instrumentation Panel in the preparation of material for use in this report:

E. I. du Pont de Nemours & Company, Inc.: H. S. Jarrett, D. R. Johnson, W. D. Phillips
The Foxboro Company: L. E. Bullock, R. P. Lawler
General Radio Company: A. P. G. Peterson
Grason-Stadler Company, Inc.: R. A. Grason
Hewlett-Packard Company: L. S. Cutler

Inland Steel Company: N. A. Robins
Instrument Society of America: F. M. Ryan
Leeds & Northrup Company: F. C. Snowden
Arthur D. Little Company: H. Boutner
Monsanto: O. R. Bailey, W. Cave, W. E. Koerner, J. R. Middleton
National Bureau of Standards: J. A. Barnes, D. B. Mann, J. A. Simpson
Scientific Apparatus Makers Association: G. Lawrence
Varian Associates: J. Feinstein, J. Holcomb, A. O. McCoubrey, D. Stangland
MacKay-Shields Economics, Inc.: F. H. Gowen

Contents

XII
INSTRUMENTATION

. . . applied technology demands a constant diet of pure research.

RONALD W. CLARK
Einstein: The Life and Times
(World Publishing Company,
New York and Cleveland, 1971)

1 Instrumentation

1.1 ITS NATURE

We live in a dimensioned world. Knowledge is projected into numbers, and precise and reliable numbers constitute the firm foundation of scientific and technological structures.

Instrumentation is the means of acquiring or establishing pertinent numbers. Thus, instrumentation may be defined as equipment that can be utilized for the sensing, acquisition, manipulation, display, and evaluation of quantitative information related to materials and energy and their processes in industry, in the physical and life sciences, and in research and testing generally. It may involve, as in automatic control systems, steps that alter quantitative data at its causative sources, so that desired numbers for certain related conditions are attained.

1.2 CLASSIFICATIONS

Instruments may be classified by a variety of matrices, and they frequently are. They may be categorized, for example, by basic elements they utilize, such as mechanical, pneumatic, hydraulic, magnetic, electrical, electronic or nuclear, or combinations of these. They may be classified by the accuracy of measurement of which they are capable. They may be identified by their area of use, such as laboratory or industrial. Or by their functions, such as indicating, recording, computing, or controlling. Or by the processes, or industries, or professions they serve, or the classes of measurements they make. Citing these many methods of classification in this report serves simply to emphasize how broad are the techniques, how widespread the uses, and how manifold the capabilities of present-day instruments and instrument systems.

Three general categories, which are not mutually exclusive, are useful for purposes of discussion:

Instruments for research and laboratory measurement and testing
Instruments for industrial process measurement and control
Analysis instrumentation, which includes types for research and laboratory use and types for industrial plant "on-line" use

1.3 THE IMPORTANCE OF INSTRUMENTATION

A society without numbers, and hence a society without instruments, is, in today's world, inconceivable. There are few areas of human activity in which instruments do not find application; some in which they do are given here.

1.3.1 *Research*

Instrumentation is a key and indispensable element in research, whether the research seeks to advance man's knowledge of nature's laws or of the characteristics of nature's materials or energies or of life and biological processes or is concerned with the synthesizing of new materials or with the development of new techniques of energy conversion or with the development of new products or processes. Each major step forward in research creates a need for new and better instruments of greater resolution or capability or flexibility, so that a next significant step forward in research can be undertaken and achieved. This general area, and the related reciprocal interplay between basic research and new instrument development will be discussed further later in this report.

1.3.2 *Industrial Processing*

In the scientific and industrial sense, ours is a well-ordered world. Fortunately, physical and chemical properties of things or materials can, by investigation, be established, and, under comparable conditions, the findings will be repeated. The whole essence of production is to produce materials and things—and, in some cases, services—with fixed, predictable qualities, qualities that can be reproduced if the conditions and environment of production are properly established and maintained.

It is the function of instrumentation first to determine the magnitude of parameters that define the conditions of a process. Then, by automatic control means, incorporating in recent times computers of great capability, appropriate variables are adjusted so that desired process conditions are achieved. Analysis instrumentation, either in the laboratory or on-line, depending on the prevailing state of the art, assists in regulating the process to yield products of desired quality while achieving economic process operation.

1.3.3 *World Competition*

Successful competition in the trade markets of the world is an essential element in maintaining a nation's high standard of living, and a correspondingly highly cultured life style. World competition is both technological and economic. Given steady trends toward equalization of technology, at least in some major fields, the competition resolves itself largely into economic comparisons, that is, comparisons of process and plant productivity.

The importance of productivity in world competition, in the face of other economic trends in the world, is illustrated in Figures XII.1 through XII.4. Figure XII.1 shows comparative wage increases over the past several years in the principal industrial nations of the "free" world. Japan, it will be seen, has had the highest rate of wage increases, with West Germany close behind. The United States has had the lowest. Figure XII.2 provides an index to changes in consumer prices. Again, Japan is the highest, with the United States between it and West Germany.

Changes in productivity as reflected in manufacturing output per manhour, are illustrated in Figure XII.3. Japan shows the greatest improvement, the United States the lowest. Figure XII.4 shows the net result of

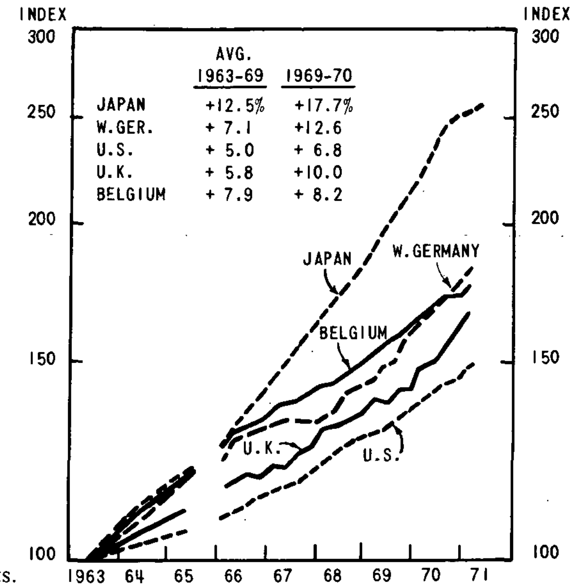


FIGURE XII.1 Wage payments.

changes in wage rates and productivity, in terms of unit labor costs. It will be seen that Japan, despite the highest rate of wage increase, has had sufficient counterbalancing improvement in productivity that its unit labor costs have been held substantially constant over the past few years.

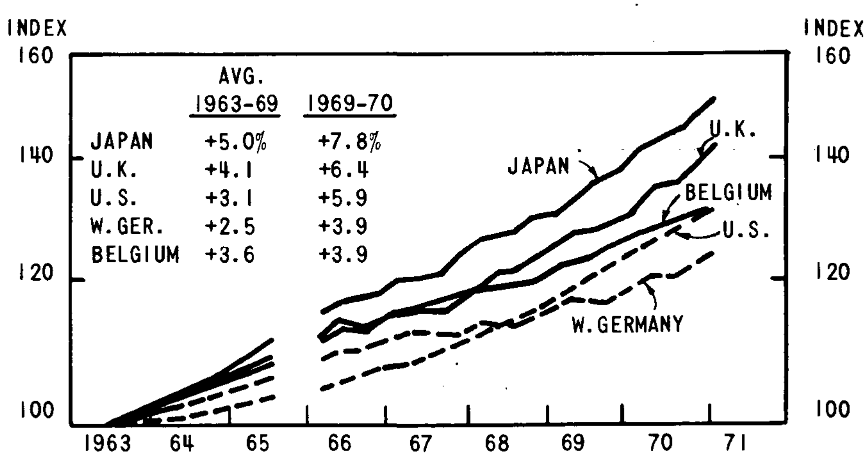


FIGURE XII.2 Total consumer prices.

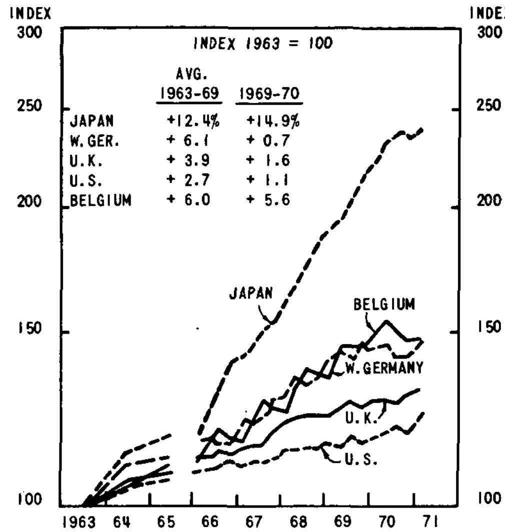


FIGURE XII.3 Changes in manufacturing output per man-hour.

The United States, along with the United Kingdom, shows the highest increase in unit labor costs.

These curves appear to state, quite emphatically, that U.S. productivity must be improved if it is to maintain a satisfactory position in world com-

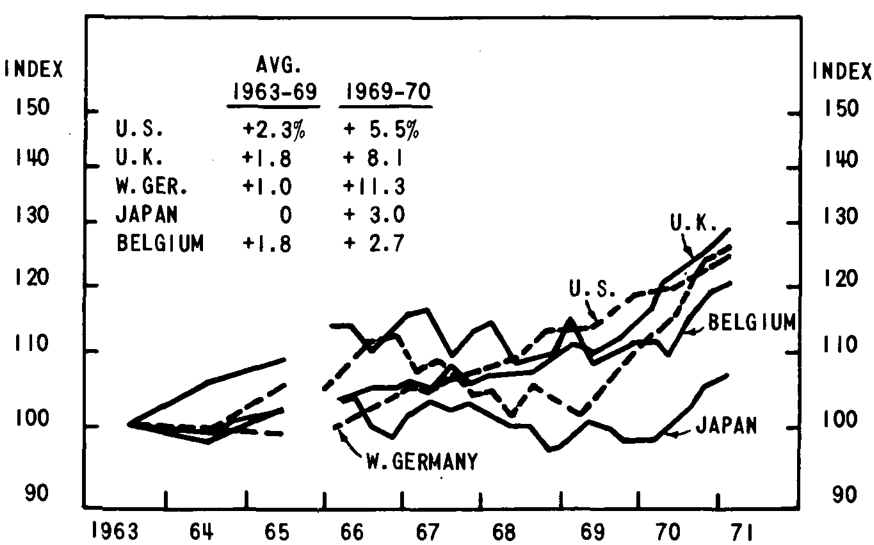


FIGURE XII.4 Changes in unit labor costs.

petition. Manifestly, the problem is not a simple one. A solution requires participation by, and significant contributions from, many sectors and groups. Advances in science and technology and related improvements in process control by means of instrumentation can be important factors in achieving the requisite productivity improvements in U.S. plants.

Meanwhile, as is discussed later in this report, the instrument industry itself continues to maintain, as it has for many years, a favorable overall balance of trade, of the order of 3.5 : 1, itself an important contribution to the nation's economic well-being.

1.3.4 *Society*

If one grants, as the Panel does, that (a) creative scientific research with its resultant contributions to the fund of human knowledge and hence to a society's cultural resources, (b) expanded and efficient industrial production that creates jobs and produces necessities and niceties for life and leisure, and (c) a position of leadership in the economy of the world to help raise living and environmental standards are desirable objectives for a nation, then, because of its pervasive use in all phases of such activities and objectives, *instrumentation can be said to contribute importantly to* society, to the health, safety, and welfare of the public, and hence to an agreeable life style.

The extent of the contributions are great, but their existence and their benefits are difficult for the public to sense or see. For one thing, they occur almost entirely within the technical community and the capital equipment industry, at stages where no consumer products or services are directly or visibly involved. For another, the benefits are diffuse, composed generally of an extremely large number of individual improvements, efficiencies, and advances in an extremely wide variety of industries, institutions, and human activities. But they are there.

As will be discussed in this report, an intimate relationship exists between scientific research and commercial instrumentation. Such instrumentation is one of the major pathways by which advances in science generally and in physics particularly are assimilated into the economy, thus achieving broad utility to society.

Thus, in speaking of the social contributions of instrumentation, one is also speaking of the social contributions of science. The Panel is well aware of the considerable criticism of science and technology that has developed in some areas. It is not intended that this point be debated in this report. It will simply be noted that the manner of use of a discipline or of products does not constitute grounds for an indictment of the discipline or the products themselves. The use of steel for swords instead of plowshares is not an indictment of steel nor of steelmaking processes.

2 The U.S. Instrument Industry

2.1 HISTORICAL BACKGROUND

During the nineteenth century, instruments used in the United States were for the most part imported from abroad, primarily from Germany and England. In the decade that followed the turn of the century, domestic instrument companies were organized and developed. Also, in 1901, the National Bureau of Standards was organized, and its pioneering work in the development of standards and measuring techniques profoundly influenced and stimulated the design, fabrication, and use of domestic precision and industrial instruments.

Thus, when World War I came, apparatus, standards, and techniques for measurement and control were available domestically to support the nation's industrial and defense needs.

In the years immediately preceding World War II, a number of small companies, largely oriented to the emerging science of electronics, entered the instrument field. They, and other new groups, supplemented the efforts of the older instrument companies in providing major support to the nation's technical effort—in both research and production—during the war. The technologies that emerged from the war, including those based on the new nuclear science and on radar activity, found expanded use in world science and industry and stimulated considerable growth of U.S. instrument enterprises. Later, space programs, expanded defense needs, extensive government-sponsored research, expansions in industrial processing, and major scientific advances in solid-state physics and computer technology all contributed to new and larger needs for instrumentation, with corresponding growth in the U.S. instrument industry and—at least until now—worldwide domination in this field.

2.2 PRESENT STATUS

2.2.1 *Size*

The current annual sales of the U.S. instrument industry are estimated at about \$3 billion and are produced in about 1800 establishments, with over 165,000 employees.

The U.S. instrument industry is characterized by a broad range of company sizes. The largest has annual sales of about \$350 million, while there are numerous companies with annual sales in the \$500,000 to \$1 million range. Small companies compete successfully with the larger ones on the basis of innovative technology and aggressive applications activity in selected product areas.

2.2.2 *Research and Development Activity*

Product lines are of "high-technology content," and while research and development activity elsewhere is an important market for the industry, the instrumentation industry itself ranks high in terms of the amount of sales reinvested in its own research and development. For individual instrument companies, the range of investment in new product development runs from 5 to 10 percent of sales. Leading companies are at the high end of this range. In 1969, the overall average for the instrument industry was 6 percent, with about 13,200 scientists and engineers employed in new product development activity at the end of that year.

2.2.3 *Contributions to Balance of Trade*

World leadership in instrumentation technology, coupled with sound business practices, has enabled the U.S. instrument industry to achieve a high level of export sales. This amounts to as much as 50 percent of total annual sales in some product areas and averages about 20 percent of total annual sales for the industry as a whole. Exports of instrument products have exceeded imports in the ratio of 3.5 : 1, thereby contributing importantly to the nation's balance of trade. Overseas instrumentation capabilities are expanding rapidly, however, particularly in Japan and Germany. Maintaining this ratio of exports to imports will be increasingly difficult in the future.

The Scientific Apparatus Makers Association, headquartered in Washington and now in its fifty-second year, is the principal national trade association for the industry. The Instrument Society of America, with headquarters in Pittsburgh and founded 27 years ago, is a professional society devoted specifically to instrumentation science and technology. It has 20,000 members and is one of the representative societies of the National Research Council.

3 New Product Development in the Instrument Industry

3.1 FUNDING

The U.S. instrument industry is a strong investor of its own funds in new product development. As already noted, the industry averages about 6 percent of its annual sales for in-house R&D, while some of the more successful companies spend 10 percent or more of annual sales in such efforts. It is emphasized that these expenditure rates are for corporate funds and do not include any special projects that may be funded within some companies by government agencies.

3.2 STEPS IN NEW PRODUCT DEVELOPMENT

Typical steps in the development of new instrument products are illustrated in Figure XII.5. In some companies, these steps occur in sequence in separate identifiable functional groups. In others, depending on each company's operating style, several of the steps shown in Figure XII.5 may be combined into project groupings. Regardless of organizational arrangements, however, the individual functions illustrated in the upper portion

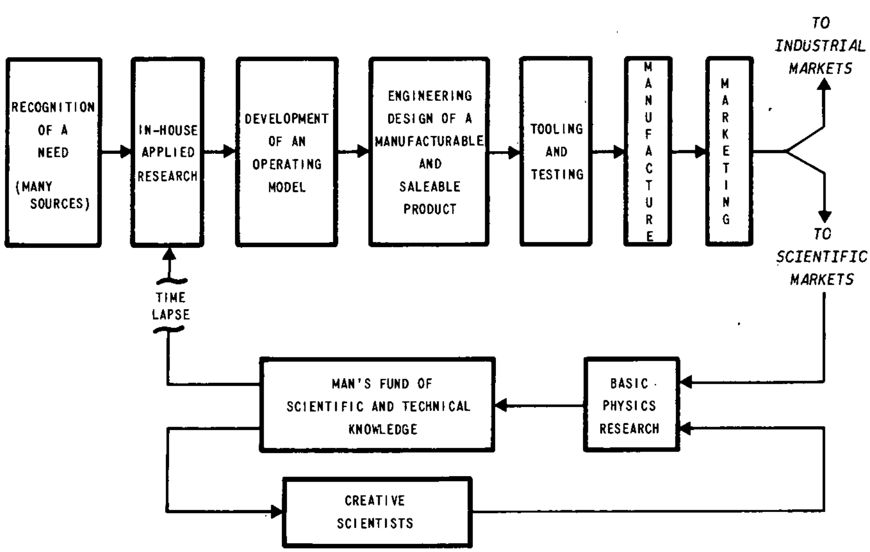


FIGURE XII.5 Steps and cycles in the development of new instruments.

of the figure—research, development, engineering design, tooling, testing, manufacture, and marketing—are the requisite elements of the overall development process, from recognition of a need to implementation in the marketplace.

3.3 TIME SPAN

The overall time for moving through the in-house steps illustrated in Figure XII.5 will vary from perhaps 2 years to 5 years, depending on the complexity of the product, the degree of innovation that it incorporates, the nature of the environments in which it will be applied, the capabilities of the purchasers' personnel, and the policies of the manufacturer concerning field failures or inadequacies of his products. Generally speaking, it can be expected that new products will have been thoroughly tested, and only after the achievement of successful results in the environments of their proposed use will they be released for general sale.

One example of a new product that required a longer time for successful development than might initially have been anticipated is the expendable immersion thermocouple for molten steel discussed in Appendix A. A relatively simple product in appearance, based fundamentally on well-known concepts of classical physics, it nevertheless required several years of intensive innovative design work and testing before fully reliable units, adapted to the rough impersonal treatment and hostile environment of a steel mill, capable of being sold for less than $1 each despite the noble metal content, were produced. The challenges to successful production and performance for such a device were great.

At the other end of the scale are highly complex analytical techniques such as nuclear magnetic resonance or electron optics, which—at least in their earlier period of introduction—are used in an atmosphere of sophistication by highly trained and sympathetic scientists or technicians, conditioned to the uncertainties of research, and possessed of a strong personal desire to have the equipment perform satisfactorily. Anticipating such an atmosphere of use may well encourage a manufacturer to cut down on the development time required for initial versions of a product destined for such markets and anxiously awaited by potential users.

3.4 RISKS IN PRODUCT DEVELOPMENT

Frequently, the time span of successful product development proves to be much longer than anticipated, and the attendant development costs much higher than originally planned. Also, it may turn out that acceptance in the marketplace may be below expectations. This may result from unanticipated economic conditions, competitive breakthroughs and offerings, or overly optimistic estimates of what potential users really would be willing to buy. All of these factors add up to risks in development, and—axiomatically—the greater the technical innovation, the greater the economic risk is likely to be. Nevertheless—equally axiomatically—the U.S. instrument industry is dedicated to taking these risks, recognizing that without them it cannot sustain growth and leadership in its selected markets.

As the industry carries out its applied research in the hope of developing profitable new products, it also hopes that in other appropriate areas there will be adequate support of basic research, first because the search for new knowledge is an intellectual challenge from which a cultured society should not shrink, and, second, because some of the results of such research will inevitably, *in time and in unanticipated ways, contribute to* the industrial growth, the economic strength, and hence the standard of living of the nation.

4 Utilizing the Results of Basic Physics Research

4.1 POSSIBLE SPINOFF FROM BASIC RESEARCH

Many new instruments incorporate the results of basic physics research, as is suggested by the feedback input to the in-house R&D blocks in Figure XII.5. Creative scientists, standing on the base of earlier discoveries and frequently using instrumentation from commercial sources (as well as creating instrumentation of their own), engage successfully in basic physics research, the results of which become significant additions to man's fund of scientific and technical knowledge. By definition, such new knowledge is sought merely for the sake of acquiring new knowledge, as part of man's insatiable curiosity about his universe and his desire to penetrate and divine nature's mysteries and secrets.

There is, in such basic research, no advance consideration of whether a practical use will be found for its results. But not infrequently, as illustrated in the diagram of Figure XII.5, and after a varying lapse of time, such use may be found, and it may prove to be a highly important use.

4.2 IN-HOUSE RESEARCH IS "APPLIED" RESEARCH

Because industry is profit-oriented, its own R&D programs have—with only rare exceptions—specific marketplace objectives, and hence are properly defined as "applied" rather than "basic" research. Basic research, in other words, is generally beyond the profit-oriented scope of instrument companies. Their R&D departments do, however, have an ongoing appetite for new knowledge. They are constantly on the lookout for new scientific information available from other sources that may in time find a place in the development of new products.

4.3 RECOGNIZING THE NEW PRODUCT NEED

As suggested in Figure XII.5, the first step, generally, in a new product program is the recognition of a market need. Sometimes such a need is clearly signaled by the demands of potential users. Sometimes it is suggested by scientists or engineers outside the company. Sometimes it is recognized by a company's field force, and sometimes by its technical staff or its management. Sometimes it is recognized as inherent in announcements elsewhere of new scientific advances, and sometimes it is but vaguely sensed as a hoped-for market response to a new product having capabilities or economic appeal beyond what had been available. Whatever the source, and they are legion, and regardless of whether the anticipated need later proves to be real or illusory, the recognition of the need is where the in-house search for supporting science and technology begins. Alert personnel will seek out every available source of supporting knowledge and of components to assist in achieving success in a new product program.

4.4 RELATIONSHIP BETWEEN IN-HOUSE RESEARCH AND BASIC PHYSICS RESEARCH

Basic research is the necessary foundation for applied technology. Expanding this structural metaphor will help place in proper perspective the utilization of basic physics in new instrument development. Given a foundation that derives from basic physics research, the architecture of the superstructure, the new product that is to find practical use, constitutes the operating objective of in-house new product development. It is the design and building of the superstructure that is the costly, time-consuming, challenging, and hazardous undertaking in the development of new instrument products.

A new product program may involve new uses of old technology and hence of old basic physics, sometimes in innovative and unusual ways, as in the development of expendable thermocouples described in Appendix A. Sometimes, on the other hand, new technology founded on relatively new basic physics is what will provide the means for achieving the desired end results.

4.5 TIME LAPSE OF COUPLING

The coupling of the results of basic physics research with in-house new product development may, as suggested by Figure XII.5, involve a varying time lapse, depending on many circumstances. The coupling may or may not be direct.

In the case of sophisticated analytical techniques such as nuclear magnetic resonance (NMR), the coupling is direct, and with a short time lapse. In other cases, the coupling insofar as the instrument industry is concerned is once removed and hence involves a longer time lag, for example, developments in solid-state physics that carry the promise of new circuit components for electronic instruments. However, ways must first be found to mass produce and manufacture these new components by companies devoted to that activity. It is their products that in turn stimulate instrument redesign and new instrument development.

4.6 EXAMPLES OF BASIC PHYSICS INFLUENCE ON NEW INSTRUMENTATION

To list all examples of instruments that have depended for their development, at least to some degree, on knowledge gained from basic physics research, is almost equivalent to providing a complete catalog of present-day instruments that are available for research, testing, product analysis, and industrial processing. Nevertheless, a tabulation is provided in this section. It includes some widely used instruments that were long ago derived from classical physics of earlier eras. It also includes newer instrumentation that is fundamentally dependent on basic physics research of much more recent times. The tabulation is as follows:

Thermocouples
Resistance thermometers
Optical pyrometers
Radiation pyrometers
Pressure-measuring instruments

Flow meters based on differential head
Electromagnetic flow meters
Magnetic resonance flow meters
Vortex precession flow meters
Level measuring meters
Precision electrical standards, bridges, potentiometers, and ratio sets
Electronic measuring and test instruments
Electronic self-balancing recorders
Time and frequency standards and instruments
Computers
Automated and computer-controlled testing systems
Automatic parts gauging systems
Automatic control systems for materials processing
Automatic control systems for energy conversion and distribution systems
Quartz crystal thermometers
Sonic sensors
Ultrasonic testing systems
Capacitance sensors
Light-emitting diodes
Solid-state components and integrated circuits
Instruments utilizing digital technology
Alphanumeric displays on cathode-ray tubes with advanced man–machine interfacing
Fluidic devices for sensing, switching, and control
Video tape stroboscopy
Thermal conductivity analyzers
Paramagnetic analyzers
Infrared analyzers
Magnetic resonance spectrometers
Neutron activation analysis
X-ray fluorescence analysis
The electron spectroscopy for chemical analysis family (ESCA)
Field ion mass spectroscopy (e.g., ion probe mass analyzers)
Scanning electron microscope
High-voltage electron microscopy ($\geqslant 1$ MeV)
Gas chromatography
High-pressure liquid chromatography
High-voltage paper chromatography
Gel-permeation chromatography
Mössbauer spectroscopy
X-ray diffraction topography

Fourier transform spectroscopy
Ion selective electrodes
High-field superconducting magnets
Solid-state detectors
Optical rotatory dispersion spectrometer
Atomic absorption spectrophotometers
Alignment lasers
Laser metrology
Laser Raman spectrometers
Microwave spectroscopy
Ion sputtering units
Low-energy electron diffraction systems
Auger analysis
Pattern recognition assemblies

The following may be added as instruments being developed but not yet in commercial use:

Tunable laser absorption spectroscopy
Molecularly stabilized lasers
Picosecond pulsing by mode-locked lasers
Josephson effect
X-ray interferometry
Holography
Echelle dispersion

4.7 SELECTED EXAMPLES FOR MORE EXTENSIVE DISCUSSION

Ten areas of instrumentation that have been influenced by, or are a direct result of, physics research have been selected for further discussion in appendixes to this chapter. Some of the appendixes discuss individual instruments or classes of instruments; others relate to fields of measurement, and one covers a complete industrial area. Another relates to future field needs. The ten appendixes are as follows: A. Expendable Immersion Thermocouples; B. Basic Temperature and Flow Sensors for Industry; C. A Walk through a Steel Mill; D. Instruments for Monitoring Environmental Quality; E. Medical and Surgical Instrumentation; F. Nuclear Magnetic Resonance Instrumentation; G. Electron Optics; H. Time and Frequency Measurements; I. Computers; and J. Some Future Needs.

5 Inflation and Costs

It has been reported that a belief frequently encountered among physicists is that costs for instrumentation used in their research work are considerably higher now than they were just a few years ago. This may well be true as an overall comparison, but for a better understanding of such views and of some of the contradictions embodied within them, it will be helpful to review the essentials of pricing and related matters in four individual categories.

5.1 WELL-ESTABLISHED INSTRUMENTS

Instruments that were already well established a few years ago, such as certain types of test instruments and recorders, successive versions of which simply reflect improvements in design details, have had inflationary price increases over the past 10 years in the range of 25 to 50 percent. Certainly, for such instruments, the impression of higher prices would be confirmed.

5.2 INNOVATIVE INSTRUMENTS

Instruments that a decade ago were relatively new in the marketplace and that reflected major steps forward in complex instrumentation, such as NMR analyzers, have—for the same performance capability—decreased steadily in price in recent years, moving significantly counter to inflationary trends. A basic assembly that sold for about $50,000 a decade ago might today sell for 60 percent of that figure. This pricing trend would seem to be in contradiction to the general impression of price increases referred to earlier. However, a research scientist today might not be satisfied to purchase just such a basic instrument, but might desire—or need—the most advanced design in order to take advantage of its greater capabilities. Such an advanced unit might cost as much as or a bit more than did the only unit available a decade ago. This advanced instrument, moreover, might be equipped with additional accessories, at still additional price, as discussed below in Section 5.4.

5.3 COMPUTERS

Computers and their constituent elements are considerably less expensive today than they were a decade ago. Figure XII.6 illustrates the drop in price in core memories from about $0.07 per bit in 1966 to less than

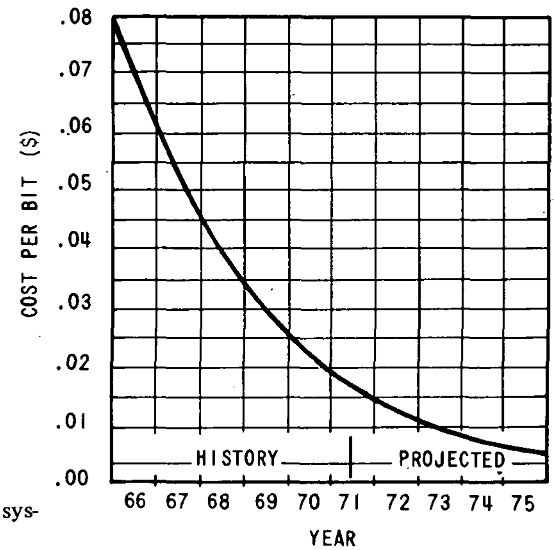


FIGURE XII.6 Core memory system cost.

$0.02 per bit in 1971. Similarly, the decrease in the price in minicomputers from about $27,000 each in 1963 to the $5000 to $7000 range in 1971 is illustrated in Figure XII.7.

In each case, these decreases reflect advances in technology of design and fabrication and substantial increases in production volume. Thus, if a research scientist were to purchase a bare computer today, he would be spending much less for it than he would have a few years ago. This, again, will seem to be in contradiction to the basic premise of increased prices expressed at the beginning of this section. On the other hand, perhaps the current substantially lower price of a computer would encourage its purchase by a scientist who might not have considered such an acquisition when prices were substantially higher.

5.4 ADVANCED SYSTEMS COMBINATIONS

It is possible that this is one of the areas that has stimulated the impression of high current equipment costs. Automated systems assemblies incorporating computer direction and computer data processing have been developed, with capabilities far beyond those available just a few years ago. It is conjectured that a research scientist, anxious to optimize his available time, to relieve himself of extensive data processing and analysis so that he can concentrate more fully on the basic scientific challenges of

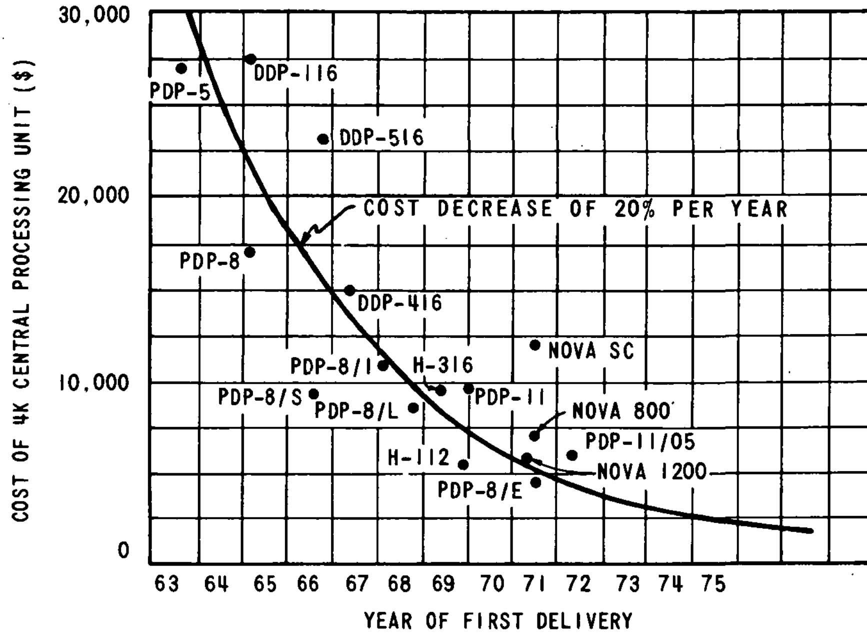


FIGURE XII.7 Decreasing costs of minicomputers.

his investigation, will seek to acquire the most advanced systems assemblies. Thus he may indeed be faced with what are truly higher equipment costs, but he may be getting a lot more than he could have acquired at lower cost a few years ago. In addition, he may be achieving secondary savings due to conservation of his own time and possibly the time of technician assistants by the automation of computational and analyses steps.

The Panel believes that taken in the aggregate, the foregoing discussion may account for general views in scientific circles that requisite investment for research equipment is now much higher than it was, even though as noted, prices on some significant units have declined.

6 Conclusion

6.1 SUPPORT OF BASIC RESEARCH

It is of interest to supplement the quotation that opened this report with another quotation from the same source that describes how basic research in physics, in another era, originally undertaken solely as an intellectual challenge for scientific discovery, later led to practical end results.

> The X rays discovered by Röntgen were giving advance warning of disease where none would have before been possible. The effect of the radium so laboriously purified and investigated by the Curies was giving the hope of life to patients who before had no hope. Einstein's explanation of the photoelectric effect had already helped to prod forward television from experiment to reality. The revolutionaries who had gathered in Brussels for the First Solvay Congress less than three decades earlier [in 1911] already had ample practical results to show for what had seemed, so recently, to be largely theoretical discussions.*

A second pertinent quotation is from testimony presented before the Subcommittee on Science, Research and Development of the Science and Astronautics Committee of the House of Representatives on July 29, 1971, by Willard M. Bright, President and Chief Executive Officer of the Kendall Company, Boston, and Chairman of the Science/Technology Committee of the National Association of Manufacturers. The subject of the Hearings was the U.S. balance of trade, which had turned unfavorable for the first time since the early nineteen hundreds. Excerpts from Dr. Bright's testimony follow:

> Industry now performs more than 70 percent of the nation's total research and development, including Federally-funded projects. To speak of industry's involvement of facilities, manpower and talent, then, is to speak of the major portion of this country's R&D effort.
>
> In broad terms, industrial R&D emphasis is directed toward applied research and development related to product and process improvement, and to new product development. It is, by and large, financed by private industry.
>
> Appropriately, the Federal Government funds the technology to satisfy the nation's military, political and social needs, and finances much of the nation's basic research. Academic and non-profit institutions are almost wholly concerned with this basic research. . . .
>
> [B]ecause successful basic research contributes greatly to the strength of the nation, it is eminently in our national interest to maintain vigorous R&D programs of basic research. The universities need to be supported in this across a broad spectrum of natural science and social science, and such support is properly given by Federal funding.

6.2 MODERN INSTRUMENTS ASSIST BASIC RESEARCH

This report has endeavored to show how basic research in physics provides knowledge that is useful in the development of new and important instru-

* R. W. Clark, *Einstein: The Life and Times* (World Publishing Co., New York and Cleveland, 1971), pp. 544–545.

ments. As suggested by the feedback loop of Figure XII.5, new instruments in turn can contribute to the effectiveness of basic research.

Comments by Chairmen of other Panels support this view:

From the Chairman of the Elementary Particle Physics Panel: ". . . almost any instrument intended for physical measurements, and even some for chemical and biological measurements, has important uses [in our field]."

From the Chairman of the Astrophysics and Relativity Panel: "Relativistic astrophysics, and indeed all of astronomy generally, is utterly and completely dependent upon instrumentation of the very highest sophistication for the detection of weak electromagnetic signals from space."

From the Chairman of the Chemistry-Physics Interface Panel: "The impact in this area is almost so great as to be beyond measure. . . . Essentially the entire field of spectroscopic instruments . . . can be traced to the field that now comprises the chemistry-physics interface. In addition to the applications stimulated by basic research . . . the demands of basic research have pushed the instrument manufacturers to develop instruments and components with far higher performance than was available, say, fifteen years ago."

The foregoing quotations provide a felicitous conclusion to this report. Basic physics has provided important foundations for instrumentation. Instrumentation, in return, has provided important tools with which scientists can extend the horizons of basic physics.

But divining nature's secrets and applying them to technologically oriented commerce and industry has no limits, no ultimates. Continued endeavor is called for, since each new achievement, however revealing or dramatic, is simply a basis for new progress. Unlike climbing a mountain, which has a clearly defined goal, scientific and technological research is never done, and whatever has become known is short of what man wants and needs to know.

It would appear therefore that the physics community and the instrument industry share related goals. It is to be hoped that, as in the past, each will in the future contribute effectively to their attainment.

> . . . "pure" science was able to move forward once again on the vehicle provided by technology. This has happened frequently. . . . Maxwell, on arriving at Aberdeen, had stated significantly: "I am happy in the knowledge of a good instrument maker. . . ." Others, commenting on the difference between the science of the twentieth and earlier centuries, have pointed out that we are on a higher level today not because we have more imagination, but because we have better instruments.
>
> RONALD W. CLARK
> *Einstein: The Life and Times*
> (World Publishing Company,
> New York and Cleveland, 1971), pp. 79-80

Appendix A: Expendable Thermocouples for the Measurement of Molten-Steel Temperatures

A.1 INTRODUCTION

To manufacture steel economically, the steelmaker must know the temperature of the molten steel in the open-hearth, basic-oxygen, or electric-arc furnace in order that the steel be poured, or "tapped," at the appropriate point. If the steel is tapped at too low a temperature, it will solidify in the receiving ladle before it can all be transferred to the ingot molds. This situation results in so-called skulls, which must be removed from the ladle, broken up, and remelted—an obviously wasteful procedure. If the steel is tapped at too high a temperature, it penetrates crevices in the mold, causing the ingot to adhere to its mold. This is called a "sticker." The mold must be broken to remove the sticker, again a costly procedure. Also, heating the steel to a temperature higher than is necessary lengthens the time of a heat and uses extra fuel, which increases the cost of the heat.

Despite intensive efforts over a period of five decades by steelmakers the world over, a fully satisfactory solution to this problem had not been developed by the mid-twentieth century. Techniques tried over the 50 years of experimentation included thermocouples of various designs and radiation-responsive units. The high temperatures and hostile environment of the measurement militated against the attainment of successful results.

In 1958, however, a thermocouple unit was introduced that, despite its noble metal content, was sufficiently low in cost to permit throwing it away after one use and highly accurate in its single use. It won rapid acceptance throughout the steel industry. It used a number of concepts derived from basic physics, although many of these concepts had been well known for decades. This device is an example of a lengthy time lapse before principles derived from basic physics were applied to practical instrumentation to fill a specific need. It emphasizes the challenges of in-house efforts to extend available knowledge effectively with innovative development and engineering activity to produce a reliable manufacturable product that can be sold at a profit.

A.2 THE BACKGROUND DISCOVERIES

Because the successful solution to the molten-steel temperature-measuring problem is a specially designed thermocouple used with a null-balance potentiometer recorder, it is of interest to sketch the historical background for these devices.

The thermocouple effect, i.e., the development of a small dc voltage related to the temperature difference between the hot and cold junctions of two dissimilar metals, was discovered by T. J. Seebeck in Germany in 1821. At just about the same time, at Berlin University, J. C. Poggendorff devised the first known null-balance type potentiometer, which he used to measure the output of electrochemical cells without drawing current from them. As far as is known, not then nor for a considerable time thereafter, was there a common interest in or joint use of the thermocouple effect and the null-balance potentiometer. It took about 80 or 90 years for the two to be combined for industrial temperature measurement.

In about 1886, some 65 years after its discovery, practical application was made of Seebeck's thermocouple effect by H. L. Le Châtalier in France, who measured the output of the thermocouple with an indicating deflection millivoltmeter. Shortly thereafter, millivoltmeter-type recorders were adapted to temperature recording using thermocouple temperature detectors. But such recorders had limitations due to insensitivity and calibration drifts of millivoltmeter movements, the influence of length of leads on the temperature measurement, and difficulties related to reference junction compensation.

Meanwhile, along another path of temperature-measurement development, the first platinum resistance thermometer, depending on the temperature coefficient of platinum, was developed in Germany in about 1871. By 1890, a practical platinum resistance thermometer was developed in Great Britain.

In 1897, the Callendar Wheatstone bridge recorder, believed to be the first null-type recording instrument on record, was introduced. It used advanced concepts of electrical balancing but did not prove too practical in the hands of other than skilled operators.

Then, early in the twentieth century, individuals in industrial plants began to combine thermocouples with manually operated null-balance potentiometers to make precise measurements of temperature. This technique, which at balance drew no current from the thermocouple, had many clear advantages and stimulated efforts to develop an automatic null-balance potentiometer recorder to take the place of the manually operated instruments. By 1912, such a successful recorder was developed, and Seebeck's thermoelectric effect came into its own for industrial temperature measurement and control. Literally, millions of thermocouples and tens of thousands of potentiometer recorders have since been applied to a wide variety of industrial processes. Nevertheless, despite the great success of this technique in other applications, it had not been applied successfully to the molten-steel temperature-measuring problem by the mid-twentieth century.

A.3 THE MEASUREMENT CHALLENGES

Some of the problems associated with making a successful measurement of molten steel were the following:

1. Temperature was at 2900°F—high for an industrial measuring device.
2. An accuracy of 5 to 10°F was desired.
3. The molten steel was covered with a surface of slag.
4. Expensive radiation-responsive devices, intended for multiple repetitive use, required heavy protection, were cumbersome, and were subject to large errors.
5. Similarly, thermocouples intended for multiple repetitive use required heavy protection and limited the operator's ability to establish temperature trends before he had to tap the furnace. Also, repetitive exposure of the platinum–platinum, rhodium thermocouple to the high temperature of the bath caused significant calibration drifts.

What clearly was needed was a device that would overcome all these limitations and permit measurements to be made rapidly and accurately at a cost not to exceed $1 to $2 per measurement.

A.4 THE SUCCESSFUL SOLUTION

The unit, introduced in 1958 that, in concert with a null-balance potentiometer recorder, successfully solved the molten-steel temperature-measuring problem is a thermocouple of very special design. Its features are these:

1. It utilizes a platinum–platinum, rhodium thermocouple, and achieves the 5 to 10°F accuracy by being used only once and then discarded.
2. It achieves the economic targets by using only a very small length of the precious thermocouple wires.
3. Base metals with suitable thermoelectric effects are used within the thermocouple in place of longer extensions of the precious wires. The junction of these base and noble metals must be less than 200°F even while the thermocouple is immersed in the 2900°F bath. This requirement posed interesting problems in regard to thermal conductivity to the thermocouple tip and thermal insulation for the noble metal–base metal junction. The combination needs to act just like a complete noble metal thermocouple for the few seconds it takes to make the measurement.

4. The tiny thermocouple tip is protected from contamination and physical damage as it is plunged through the furnace atmosphere and the surface slag.
5. The thermocouple holder, or lance, is light enough to be handled easily by one man and yet heavy enough to pierce the slag and metal surfaces and withstand general rough steel-mill use.

Extensive development and design work, tooling, and testing ultimately produced the present-day expendable devices that sell for considerably less than \$1 each. In the 13 years since their introduction, over 100 million units have been used in steel mills around the world. This was a case, then, of the basic foundation being available from classical physics; the challenging applications problems awaited imaginative, innovative design work before they were resolved.

Appendix B: Basic Temperature and Flow Sensors for Industry

B.1 INTRODUCTION

Temperature and flow rate are fundamental measurements in almost all industrial plants and in commercial custody exchange points. Two 1971 symposia,* jointly sponsored by the National Bureau of Standards, the American Institute of Physics, and the Instrument Society of America, summarized a decade of progress in sensors that measure these two key variables. Papers in both symposia demonstrated that progress in science and progress in the basic sensors have been interdependent.

Temperature measurement, because of its convenience, has become a nearly universal inferential measurement of other variables whose direct measurement is less convenient. For example, in a flow-calibration laboratory, one infers viscosity by coupling the measurement of fluid temperature with look-up tables. In polymerization processes, temperature controls stabilize the chemical kinetics, although the kinetics are affected by many other variables.

* Flow Symposium, May 10–14, 1971, Pittsburgh, Pa.; Temperature Symposium, June 21–24, 1971, Washington, D.C.

While the fidelity of sensors for measurements of temperature and flow rate is strongly dependent on installation geometry, the dependence is especially inherent in sensors of flow rate. This is because the flow sensor, in principle, integrates all velocity vectors across the flow-measurement section. If the velocity profile is fixed, as it is in precision tests of power turbines, the consistency of the integration is high and flow measurement may be good to ±0.1% of range. But in plant applications, upstream piping geometry that is usually far from ideal can skew the velocity profile, seriously downgrading the consistency of integration, and, therefore, downgrading the predictability of sensor operation unless the entire installation is calibrated.

B.2 IMPACT ON PRODUCTIVITY

A distillation tower in a petroleum refinery and an evaporator that concentrates digester liquor in a pulp mill depend upon management of thermal energy transfers (temperature measurements) and liquid inventories (flow and level measurements). Computer management of the transfers, and inventories within the physical limitations of the apparatus, increases the efficiency of the use of energy and feed materials. While the increase in efficiency may be reported in only a few percentage points, the reduction in the ratio of operating costs to production volume is significant. This method of apparatus management, expanded with on-line measurements or calculation of product quality, is behind the trend for computer control in industrial plants. For the trend to gain momentum, *in situ* determination of process characteristics is necessary. Therefore, the computer and the associated flow and temperature sensors become an experimental system for process engineers and an apparatus management system. The dual-purpose system creates insistent demand for flow and temperature sensors with wider range and accuracy.

B.3 IMPACT ON INTERNATIONAL BALANCE OF TRADE

U.S. manufacturers find ready export markets for flow- and temperature-sensor systems. The ratio of export to import favors the U.S. manufacturer by greater than 3 : 1. Technological advantage in the sensor, its application procedures, and its service at point of delivery account for favorable balance. Other countries have the technical knowledge but have not yet developed it into a business. Because they will, the U.S. manufacturers must respond to the demand for sensors with wider range and accuracy.

Manufacturers must also respond to the drive for an international system of measurement units (SI). Mass, length, time, and temperature, four of the six basic SI units, define the reference standards and the markings of flow rate and temperature sensors.

B.4 BENEFITS DERIVED FROM SENSOR DEVELOPMENT

With control of temperatures and flows so vital in almost all industrial processes and energy-conversion systems, effective sensors contribute to plant productivity, operations economy, and product quality, which, in turn, are important elements in making attractively priced products of merit available to the general public.

Because the flow of water and every fuel is metered for transportation and consumption, flow sensors are common but unheralded monitors of fluids that directly affect the quality of life. Many of the everyday applications of temperature measurement in the household, in reporting the conditions of our natural environment, and in regulating our artificial environments are already well known. However, new applications, such as measuring the quantities of effluents from smokestacks and plants and directly measuring thermal pollution, loom large in the public eye.

In other examples, marine ecologists map thermal energy currents and find that the rate of change of thermal energy may be deadlier to marine life than is high temperature itself. Flow-rate and temperature-rate measurements thus combine in marine ecology to produce a single, more significant, measurement of the quality of the environment for marine life. Long-wavelength cooled infrared detectors map thermal abnormalities in the human body to seek malignancies. Differential thermal analyzers are used in research on the DNA molecule.

B.5 SOME PRIORITIES OF APPLIED RESEARCH

In flow sensors the most promising work centers around: (1) Improving the predictability of the widely used head meter. The Fluid Meters Research Committee of the American Society of Mechanical Engineers coordinates this work nationally by voluntary effort and financing. Research is necessary to cleanse the data of early experiments that were poorly controlled and to establish data for predicting installations with the nonideal upstream conditions that prevail in industry. (2) Developing the accuracy of wide-range hydrodynamic oscillators (the vortex meter is one form) and tagging methods. The physics of fluids and applicable electronic data-processing techniques must be investigated scientifically. Engineering empiricism has shown the feasibility. (3) Developing sensors that measure actual, not inferred, variables. Gas distribution utilities need to sense therms, not just compensated flow rate or even mass flow rate. (4) Developing signal-processing techniques, now practical because of large-scale integration (LSI) components. Instead of adulterating the measured fluid with a thermal, chemical, or radioactive tag, compact autocorrelators may measure the velocity of inherent turbulence tags of flowing fluid.

In temperature sensors: (1) Work on new materials for thermocouples with reduced drift and high-temperature life and resistance thermometer detectors with lower resistance and cryogenic ranges will, it is hoped, move forward under scientific control. (2) Certain pyroelectric materials show promise for measuring rate of change of temperature. Noise thermometry, perhaps because of its independence of material, appears attractive for ranges from cryogenic temperatures to 700°C. (3) The dependence of radiation thermometry on apparent temperature may yield to variable electronic processing with set optics. (4) The recent temperature symposium demonstrated interest in resonant techniques. Of these, resonance by inductance change is limited to below room temperatures; microwave cavity resonance is subject to geometric instability and poor signal-to-noise ratio; resonant quartz has excellent characteristics but must compete in the temperature ranges dominated by inexpensive thermocouples; but metals and ceramics at ultrasonic frequencies merit extensive investigation.

Applied research on temperature sensors benefits from fundamental research on calibration standards. The calibration background continually enriches the sensor foreground. One powerful result, for example, is that a figure of merit exists for across-the-board statements of noise. Noise is expressed as an equivalent Kelvin change. The field of flow sensors could be better ordered by such an absolute figure of merit. One finds it in magnetic-flow metering but not in head metering, the workhorse flow sensor.

Appendix C: A Walk through a Steel Mill

C.1 INTRODUCTION

C.1.1 *The Nature of Industrial Processing*

Industrial processing is primarily a matter of altering nature's materials in a sequence of steps, each carried out under controlled conditions and environments, to yield useful intermediate or final products. Instrumentation is the means of monitoring and controlling conditions at each processing step, so that desired end results of predictable nature, character, and quality may be obtained.

A modern integrated steel mill is an example of such multiple-step processing that makes use of a broad range of physics-derived sensors, signal processors, display devices, computers, and controllers for automatic regulation of successive operating steps. This appendix presents an essay tour through such a steel mill and shows what the nature of its major process steps are and how instrumentation and computer control are utilized. The physics origins of much of the instrumentation that is described will be readily recognizable.

C.1.2 *Steel Manufacture*

The manufacture of steel consists of a series of individual processes ranging from the high-temperature reduction of iron ore to molten iron to the low-temperature mechanical rolling of steel sheet for flatness and strength. Computer process control has been applied with varying degrees of success to virtually every process in the sequence. Instrumentation, of course, is the basic ingredient in any computer control system, and it is of interest to consider the various kinds of instruments that currently are used in steel-industry computer-control systems.

For convenience in presentation, several specific processes are considered in detail. For each process, a general description of the process and the operation of an associated computer-control system is given first; then, the instrumentation involved in each computer-control system is broken down into those that measure some property of the steel in process and those that measure some property of the process itself. Finally, some attention is given to specialized instrumentation used in computer-control systems on other processes, without going into detail on all the instrumentation used for those processes.

C.2 BLAST FURNACE

C.2.1 *General Description*

The blast furnace is a large column in which iron oxides, coke, and limestone are charged at the top and hot air is introduced near the bottom. The air reacts with the coke to produce the gas required to reduce the iron oxides to iron and the heat required to melt the iron and the lime. The molten iron and slag produced collect in a pool at the bottom of the column and are removed periodically. The hot air at the bottom and the solid materials at the top are charged continuously.

Process control of the blast-furnace process consists, in general, of regulating the process to produce a maximum amount of iron in a minimum amount of time with a minimum amount of coke. To this end, the computer controls the makeup of the raw materials charged to the top. This includes the relative amounts of different iron-bearing materials, the ratio of iron-bearing materials to coke, and the ratio of iron-bearing materials to limestone. Also controlled are the makeup and flow rate of the blast, i.e., the hot air introduced at the bottom of the column, which may include moisture or auxiliary fuels as injectants. Production-rate increases of 4 percent and coke rate decreases of 1 percent have already been attributed to the computer control system.

C.2.2 *Instrumentation*

The solid materials are lifted to the top of the furnace in a skip. Direct measurements of the drive-motor current and voltage and a measurement of the skip velocity with a pulse tachometer are combined to given an indication of the weight of material in each skip. Analysis of the material is performed off-line by a vacuum spectrograph or by wet chemistry, although some in the industry have attempted to use on-line x-ray analysis. The level of material in the column can be measured with either a beta-ray or photocell device. The moisture content of the raw materials charged can be measured by a nuclear device in which the absorption of gamma radiation is proportional to the density of the material and the absorption of neutrons is proportional to the hydrogen present. In general, no further measurements are made on the iron-bearing material until several hours later when it runs molten out of the furnace during tapping. At this time, the temperature is measured by an expendable immersion thermocouple or by a total radiation pyrometer, and samples are taken for off-line analysis in a vacuum spectrograph.

With regard to other aspects of the process, the flow rate, temperature, and pressure of the hot blast are measured by orifice meter, thermocouple, and pressure-to-current transducer, respectively. The moisture content of the hot blast is measured by dew cell, and the amount of fuel injected is measured by orifice meter. The flow rate of gas leaving the top of the column is measured by an averaging pitot tube. The carbon monoxide and carbon dioxide contents of the top gas are measured by infrared analyzers, and the hydrogen content is measured by a thermal conductivity cell. The top-gas pressure measurement is the same as that for the hot blast. Furnace-cooling-water flow rates are measured by orifice meters, and the temperatures are measured by resistance bulb thermometers.

C.3 BASIC OXYGEN FURNACE PROCESS

C.3.1 *General Description*

The basic oxygen furnace process is a batch steelmaking process in which carefully weighed amounts of hot metal, scrap, and flux are charged to a tilting furnace and blown with pure oxygen. The oxygen reacts with the impurities in the hot metal, thus removing them and providing sufficient heat to melt the scrap. After about 15 min, the oxygen blowing is stopped, and, if the composition and temperature of the steel are proper, the steel is poured out of the reactor into a ladle.

In this process, the object of a computer-control system is to ensure the production of steel of the proper temperature and composition in the least time with the minimum use of oxygen and raw materials and the minimum lining wear. Before the blow, the computer calculates and displays the required charge weights. During the blow, the carbon content of the steel is calculated and displayed as a guide to the operator as to when to stop the blow. In some systems, the stopping time, as well as the blowing rate and position of the oxygen lance, are controlled by the computer. Percentage increases in productivity due to computer control are small, but the high throughput rate of the process makes the rate of return attractive.

C.3.2 *Instrumentation*

The main component of the weighing scales are load cells, either hydraulic or mechanical. Composition of hot metal is measured off-line by vacuum spectrograph. Measurement of oxygen flow rate, temperature, and pressure is standard. Oxygen lance position is measured by a digital pulse-generating system. Pitot tubes, Venturi meters, and simple pressure taps have all been used to measure the flow rate of the exhaust gas from the furnace. A variety of methods is used for exhaust gas analysis. As in the blast furnace, carbon monoxide and carbon dioxide may be measured by infrared absorption. Oxygen can be measured by a paramagnetic analyzer. Alternatively, the analysis of all the gases in the exhaust gas, including nitrogen, can be obtained from an on-line mass spectrometer. Chromatographs have also been applied in this application, as well as in the blast furnace.

Because of the inaccessibility of the bath during the oxygen blow, a variety of techniques has been developed by different operators for making external measurements that can be related to what is happening internally. For example, a calibrated microphone can be used to measure the sound level during oxygen blowing. A total radiation pyrometer aimed at the mouth of the vessel can be used to indicate the completion of carbon removal. As a final example, in one system strain gauges mounted on the sides of the furnace are used to indicate the internal activity.

A variety of auxiliary measurements such as cooling-water flow rates, temperatures and pressures, vessel position, and various status switches are used in most process-control systems.

C.4 HOT STRIP MILLS

C.4.1 *General Description*

In the hot strip mill, cold slabs of steel are reheated to rolling temperature, rolled at high speed through a multitude of rolling stands to approximately 1/100 of the thickness of the slab, cooled on a runout table, and coiled. The object of most hot-strip-mill computer-control systems is to maximize the throughput of steel while maintaining close metallurgical and dimensional controls—all to be accomplished with minimum total power and within the maximum power restrictions of each mill stand.

The focal point of the control system is the rolling mill itself. A computer sets the unloaded roll openings and stand speeds and controls the feeding of the slab to the mill. While the slab is in the mill, the computer accelerates the mill to maintain temperature control and varies the roll openings to maintain dimensional control. As the tail end of the slab leaves each stand, the computer immediately resets the stand speed and unloaded roll opening for the next slab. Typically, such a computer-control system accounts for productivity increases of 5 to 10 percent.

C.4.2 *Instrumentation*

Following the steel through the process, the slab furnace temperature in the reheating furnaces is measured by a total radiation pyrometer. During its passage through the mill, the slab thickness is measured at several points by x-ray gauges and the temperature by total radiation pyrometers. For the latter, narrow-wavelength pyrometers and two-color, or ratio, pyrometers have also been used. Infrared hot-metal detectors are used to indicate slab location in the mill. Also, an infrared gauge is used to measure strip width.

Returning to the slab reheating furnace, fuel-flow and airflow rates are measured by orifice meters, zone temperatures are measured by thermocouples or total radiation pyrometers, unburned exhaust gas by combustion or Btu meters, exhaust-gas oxygen content by paramagnetic oxygen analyzers, and, in some cases, exhaust-gas analysis by infrared or thermal conductivity analyzers. In the mill itself, stand power is obtained from measurements of current and voltage directly; roll force at each stand is obtained from hydraulic or mechanical load cells; unloaded roll openings are obtained from digital position encoders measuring screw-down position; and stand speeds are measured by pulse tachometers. A variety of switches–photoelectric, magnetic, pressure, and mechanical–are used to indicate on–off or positional status of various components of the mill.

C.5 COLD ROLLING MILLS

C.5.1 *General Description*

In a cold rolling mill, steel at ambient temperature is played off a coil, passed through a multistand mill where its thickness is reduced up to 80 percent, and recoiled. As in the hot strip mill, the objective of the control system is to obtain maximum throughput with minimum energy while maintaining dimensional tolerances and stand horsepower limits. However, in contrast to the hot strip mill, temperature is relatively unimportant, while flatness and surface appearance are extremely important.

In this process, a computer also sets the unloaded roll openings and stand speeds prior to the start of rolling. During rolling, however, the computer primarily controls stand speed in order to vary the tension between stands to compensate for minute variations of thickness from a set point. In the event that tension limits are reached, unloaded roll openings can also be changed by computer. The primary benefit of the computer-control system is closer dimensional tolerances on the finished product, which enhances the competitive position of the company with its customers.

C.5.2 *Instrumentation*

The major measurements made on steel during cold rolling are the thickness and tension between each stand–the former by x-ray gauge and the latter by hydraulic or mechanical load cells. Also, overall strip length is measured by digital pulse counter and coil diameter by a photocell device. Several instruments have been used to measure flatness, although none is universally accepted. In two of the devices, a series of rollers spaced across the width of the strip is used. In one of the systems, the deflection of the rollers relative to each other is measured, and, in the other, the force exerted upon each roller relative to the others is measured. Another type of flatness gauge uses eddy-current devices to measure the magnetic permeability across the width of the strip. This is related to the tension across the width of the strip, which, when the strip is under sufficient overall tension, can be directly related to flatness.

On each stand of the mill, measurements of roll force, unloaded roll opening, horsepower, motor current, and stand speed are made in a manner similar to that of the hot strip mill. In addition, roll bending forces are measured by load cells, and the rpm of the coiler reel is measured by a digital pulse counter for automatic slowdown control.

Although not yet directly incorporated into process-control systems, considerable interest is evident currently in devices to detect surface or internal defects in the strip. Internal defects and some larger surface defects can be detected either ultrasonically or by the use of eddy-current techniques. Surface defects can be detected by photocells or video cameras under proper illumination such as fluorescent, sodium vapor, mercury vapor, or laser light.

C.6 GENERAL

C.6.1 *Other Processes and Instruments*

In addition to those processes covered above, several other processes in the steel industry are being computer controlled. Among these are electric furnaces, slabbing mills, billet mills, sinter plants, heat-treating lines, and merchant mills. Rather than cover these processes in detail as was done for those already described, only instrumentation that is different from that mentioned above will be covered here.

In the sinter plant, moisture content of the bed can be determined by measuring its electrical conductivity with two probes. A similar device is used in the blast furnace to ensure the proper filling of skips with coke. In the electric furnace, power is measured by a pulse counter attached to a wattmeter. On a continuous annealing line, hardness of the steel can be measured continuously by a device that impresses a magnetic spot on the strip and then measures its retention some distance downstream. Slabbing mills and structural mills are more or less similar to hot strip mills as far as instrumentation is concerned. However, on merchant mills, instruments to measure bar diameter and roundness are being tested. Such instruments are generally optical in nature, using either lasers or lens systems to provide collinear light; but there is at least one system that uses the infrared radiation of the hot bar. In all of them, however, the fraction of the total emitted light picked up by photocell or infrared detector is acted upon by a series of electronic components to produce the desired signal. In zinc coating lines, coating thickness is a major consideration, and this measurement is generally obtained by an x-ray fluorescence instrument.

While the emphasis of this section has been on primary sensing elements, it should be noted that no computer-process-control system could exist without the large family of conversion devices generally available. Computers are capable of taking only electrical signals, and a variety of transducers and retransmitters are required to provide the proper signals from the outputs of the primary sensing elements. Furthermore, power supplies, electronic filters, and various solid-state devices are integral parts of all computer-control systems, without which no computer could "see" the process that it is controlling.

Appendix D: Instruments for Monitoring Environmental Quality

D.1 INTRODUCTION

The present phase of scientific and technical activity related to problems of environmental quality is best characterized as an analytical phase. Even though numerous pollution control and abatement measures are being actively instituted and enforced, these actions are directed at situations that have been recognized for many years and are based primarily on approaches that were developed in an earlier era. Only recently has the all-pervasive, intricate, and critical nature of environmental problems become apparent, and even so it has become apparent only qualitatively. The quantitative data and thorough understanding required for a truly effective campaign against pollution are lacking. Accordingly, a major effort must be directed toward determining and understanding where dangers lurk, how serious they are, and what mechanisms are involved. As deeper understanding is acquired, more effort can and will be profitably directed to means of solution. Physics will surely contribute to these solutions, and it may be expected to spawn entirely new approaches to industrial activity, some of which will simply bypass older methods and their deleterious side effects. The quest for means of generating electrical power utilizing the "clean" nuclear fusion reaction is one such effort.

In the meantime, the contributions of physics in the analytical phase of the attack on pollution are numerous, as will be apparent from the examples that follow. It will be recognized that many of these contributions consist of providing advanced instrumentation for analytical chemists, who will doubtless play a major detection and monitoring role in the immediate future.

D.2 ANALYTICAL TECHNIQUES FOR AIR AND WATER POLLUTANTS

Some typical examples of analytical techniques that are being used in instruments for environmental monitoring of such pollutant sources as stationary stacks, reactors, and industrial and municipal outfalls are the following:

D.2.1 *Tyndall Effect*

This basic phenomenon occurs because suspensions of extremely small particles in water (colloids) or in air (aerosols) scatter light, which can thus be detected and measured by placement of suitable sensors at some angle to the main beam of light.

Instruments based on this principle include turbidimeters of various designs for measurement of suspended solids in water or wastewater, nephelometers for similar detection of nonsettleable particles or droplets in the air, and newly developed devices using the measurement of laser backscatter for detection of suspensions in both water and air.

D.2.2 *Surface Reflectance*

Light impinging upon surfaces, liquid as well as solid, is reflected back in an altered condition, which is a function of the characteristics of the surface. On this principle rests the design of such instruments as reflectometers for determining oil films on water, tape samplers for the continuous batch measurement of particulate matter, and a fairly recently developed reflectance turbidimeter.

D.2.3 *Light Extinction*

This occurs when the intensity of a light beam is cut back by the presence of particle suspensions in the beam. The basic concept of smoke samplers of various designs, Ringlemann number charts, and simple turbidimeters are among the devices dependent on light extinction for their operation.

D.2.4 *Diffusion-Selective Membranes*

Diffusion of gases, both in solution and in the gas phase, proceeds selectively through certain membranes. This process is a function both of membrane pore size and the adsorption–desorption characteristics of the various components of a gas mixture on the material of the membrane. Examples include the use of membranes to separate O_2 from other gases in solution to provide a polarographic measurement of dissolved oxygen. Similarly, selective membranes are used to control diffusion of desired gases into a fuel cell, where, by controlled reactions, an electrical current proportional to the concentration of the diffused gas in the cell occurs. Instruments based on fuel-cell technology are available for both ambient and source monitoring of air pollution.

D.2.5 *Heat Conduction*

All gases conduct heat at a characteristic rate; this is one of the older recognized physical phenomena on which analytical methodology is based. Thermal conductivity analyzers are designed to take advantage of this phenomenon. The method is useful for measuring SO_2 in percentage concentrations in such applications as ore roasting, copper smelting, and sulfur burning.

D.2.6 *Thermal Coefficient of Resistivity*

In another long-recognized physical principle the electrical resistance of most metals increases with increasing temperature. This principle underlies the functioning of hot-wire (platinum, tungsten, etc.) detectors used in thermal conductivity analyzers, thermomagnetic oxygen analyzers, gas chromatographs, and others that have found utility in measurement of gaseous components in stationary air-pollution sources.

D.2.7 *Paramagnetism*

Paramagnetism is a physical property that is nearly unique with O_2 among gases (two of the nitrogen oxides and a few other gases exhibit slight paramagnetism). Analyzers based on pure paramagnetic measurements, as well as in combination with paramagnetically induced convection currents (thermomagnetic) are available for specific analysis of O_2 in stack gases. This is becoming an increasingly important measurement in pollution abatement applications, both for control of low-excess air firing in combustion of fossil fuels and to provide evidence of nondilution of combustion gases in stacks.

D.2.8 *Ionization of Solids in Solution*

This is a basic and widely applied physical or physical-chemical principle, first recognized and explained by Arrhenius. Most inorganic, and many organic, compounds ionize when placed in aqueous solution, i.e., they break up, or become dissociated, into electrically charged particles, which in turn cause the solution to become electrically conductive. The analytical techniques and instruments based on this phenomenon are literally legion. Among them are measurement of electrolytic conductivity, hydrogen ion concentration (pH), selective ion measurement, oxidation/reduction potential (ORP or redox), polarography, and coulometry. Techniques such as dialysis, electrodialysis, and reverse osmosis are based, in part, on the phenomenon of ionization.

D.2.9 *Radiant Energy Absorption*

Most materials absorb radiant energy at characteristic wavelengths in some region of the electromagnetic spectrum. This basic physical principle also forms the basis for a wide variety of instruments used in environmental measurements. Examples of such instruments are dispersive and nondispersive infrared and ultraviolet analyzers used to determine concentrations of pollutants, e.g., SO_2, NO_2, and CO, in stationary sources and automobile exhausts, as well as in the ambient air. Radiation absorption is also basic to such instruments as colorimeters, spectrographs, and spectrophotometers, all of which have been used in the laboratory for water- and air-quality determinations and some of which have also been employed in the field. See Sec. D.4 for additional comments on such techniques.

D.2.10 *External Stimulation of Coherent Radiation*

A relatively recently discovered physical characteristic of certain materials, e.g., ruby crystals and rare-gas-filled tubes, is the generation of coherent radiation within the material by external stimulus of short bursts of energy. The resultant lasers thus provide a unique source of radiation of very narrow wavelength, which can be used with the instruments described in Secs. D.2.1 and D.2.9. Recent developments using tuned lasers as sources of infrared radiation are currently being investigated by the National Bureau of Standards under contract from the Environmental Protection

Agency for use in measurements of SO_2, NO_2, CO, and other gases in stationary sources.

D.2.11 *Adsorption–Desorption*

The differing adsorption–desorption properties of gaseous compounds on the surface of pulverized solids, which may or may not be coated with selective, high-boiling compounds, provides the basic principle behind all gas-phase chromatography. This method of analysis, with suitably sensitive detectors, is finding wide application in both ambient and source air-pollution monitoring.

D.2.12 *Radiation Emitted by Heated Gases*

Many gases, when heated, emit radiation at wavelengths that are characteristic of that gas. This has resulted in development of flame photometers that can provide reasonably selective measurements for sulfur gases, as an example.

D.2.13 *Ionization of Gases at High Temperatures*

Many gases, oxidizers for the most part, will ionize at elevated temperatures. This principle is used in the flame ionization detector (FID) by feeding the sample, under closely controlled conditions, into a stream of hydrogen gas that is equally well controlled. Upon burning between two charged electrodes, the sample gas ionizes, causing a measurable change in the electrical conduction through the flame, which can be related to its concentration. The FID provides a sensitive detector for the binary mixtures that elute from gas chromatographic columns.

D.2.14 *Other*

Other basic physical phenomena are under investigation for possible application to environmental analytical instrumentation. As examples, chemiluminescence may provide a way to measure extremely small concentrations of ozone and other oxidizing gases; microwave spectrometry (based on the principles discussed in Sec. D.2.9) has been suggested as a possible way to measure gases in stacks.

D.3 NOISE POLLUTION

Noise is an increasingly important problem in our highly industrialized civilization. Aircraft, trucks, powerful construction equipment, and high-speed production machinery are ever-present examples of the many intense sources of noise. Isolation from these noise sources is made ever more difficult by the increased use of lightweight and open construction.

Many disciplines contribute to the study and control of noise. Experts in psychophysics and biophysics examine the reaction of man to noise and the effects of noise on hearing loss and work output. Physicists and engineers develop noise-measurement techniques and instruments for assessing noise and for tracking down and understanding the phenomena that produce noise. They develop noise-control techniques that reduce noise at the source, and they develop quieter ways of doing required tasks.

The principal tool in the control of noise is instrumentation. The sound-level meter is the most common of the instruments used. The development over a period of many years of better microphones, circuits, and components has led to improved accuracy and reliability and to small, convenient measuring devices, for monitoring noise from trucks, automobiles, and motorcycles; to assess neighborhood noise; and to monitor noises that may be potentially harmful to hearing. In addition, one of the critical areas now being studied is the subjective effects of impulse noise and the instrumentation required for such studies. Impulse noises are common; for example, the sounds produced by typewriters and punch presses are in this class.

When noise control is the prime consideration, detailed analysis is often essential in tracking down the source of the noise and in reducing it economically and efficiently. Wave analyzers have been used for this purpose for many years. The recent development of time-compression analyzers and dedicated computer systems are revolutionizing these techniques.

New monitoring procedures and ratings for airport-noise control are being developed. The large number of operations involved requires that the measured noise values be fed into a computer, which processes the data from many monitoring points with respect to level and time exposure to yield exposure indices. These may be correlated with individual aircraft takeoffs and landings in order to effect detailed control of the operations of the aircraft with respect to noise.

D.4 FUTURE NEEDS

It is important to emphasize that much research work remains to be done before fully reliable and precise results using the techniques discussed in Sec. D.2 and similar ones for air and water pollution are routinely available. Most of the available methods require careful control of critical conditions, components, or accessories. Considerable improvement and innovation will be necessary before repetitive results, without the use of highly skilled personnel and consistent with the requirements of effective antipollution standards, can be obtained.

Also, substantial needs exist for highly sensitive measuring techniques for trace quantities of specific pollutants in natural water. Continuous monitoring instrumentation is especially needed for the measurement of pesticides; phenol; nutrients, including phosphates and nitrate; trace quantities of heavy metals, especially mercury, lead, cadmium, zinc, and copper; and the determination of microorganisms, especially coliform bacteria. Presently available instrumented measurement techniques for these parameters are restricted to well-equipped laboratories, and field monitoring instrumentation is still in the research stage.

Promising new techniques, based on advanced physical research, some of which have already had initial uses in environmental monitoring and all of which may in the future play an important role in instrumentation for this field are the following:

D.4.1 *Molecular Correlation Spectrometry*

In 1964, molecular correlation spectrometry was developed as a technique to increase the sensitivity of astronomical studies of the light absorption by water vapor in the atmosphere of Venus. The technique has been used in monitoring pollutant gases such as SO_2 and oxides of nitrogen and in mapping pollutant concentrations over large metropolitan areas.

D.4.2 *Optoacoustic Spectroscopy*

Optoacoustic spectroscopy makes use of a recently developed tunable spin-flip Raman laser to shine monochromatic light through a gas absorption cell. The light is periodically interrupted, and, hence, heating of the gas from absorption of the light is periodic and results in an acoustic signal that is detected with a sensitive microphone. Optoacoustic spectroscopy has had some use in detecting nitric oxide (NO), a major constituent of pollution from electric power-generating plants and from internal combustion engines.

D.4.3 *Tunable Semiconductor Diode Laser*

Tunable semiconductor diode lasers have recently been developed that allow infrared spectroscopic measurements to be made with a resolution about 10,000 times greater than can be achieved with conventional infrared grating spectrometers. This development may find use in point sampling of pollutants and in remote detection of pollutants present in smokestack effluents.

D.4.4 *Remote Sensing of Air Pollutants by Laser Radar Techniques*

Laser radar (or lidar) is in some respects similar to radar but differs in several important ways. A laser, rather than a microwave transmitter, is used as the source of pulsed electromagnetic radiation. Another important difference is that by virtue of the Raman effect, laser radiation scattered by some gases may be shifted in wavelength from that of the incident radiation by an amount that is characteristic of the scattering gas. Hence measurement of this shift identifies the scatterer. Laser radar, though still in its infancy, shows promise as a means of mapping distributions of Raman-active gaseous air pollutants.

D.4.5 *Neutron Activation Analysis*

Neutron activation analysis, a development from nuclear-physics research, is a highly useful analytical technique for detecting trace impurities and has exceptionally high sensitivity. Originally used with great success for analysis of high-purity semiconductors in the electronics industry, neutron activation analysis is now beginning to contribute to the solution of environmental pollution problems. The method has been useful in detecting trace amounts of mercury in fish and birds and in identifying the sources of oil spills in the oceans. Oils from different sources contain different quantities of trace elements, which are readily identified with the help of neutron activation techniques. The trace-element compositions of some 200 oils have already been cataloged as part of a library that will assist in identifying offenders in oil-spill incidents.

D.4.6 *Ion Microprobe*

In the ion microprobe, a finely focused continuously adjustable ion beam impinges on a selected microscopic region of a sample, vaporizing and ionizing atoms from the surface of the specimen. These ions are collected and analyzed with a mass spectrometer, which is attached directly to the microprobe. This system is sensitive to all the chemical elements and will identify their isotopes as well. It can detect 10 parts per billion in a 10^{-18} gram sample, i.e., a particle having linear dimensions of the order of a millionth of an inch. The ion microprobe is ideally suited for the analysis of both organic and inorganic particulate matter from polluted air and water. Because of its ability to distinguish among a variety of particles in a collection, the ion microprobe may pinpoint the origins and their relative contributions to a multiple-source particulate pollution problem.

D.4.7 *X-Ray Fluorescence Analysis*

X-ray fluorescence spectroscopy is a relatively old but effective analytical technique that has recently attained greatly enhanced capability through use of much improved lithium-drifted semiconductor detectors and associated systems. An interesting application of x-ray fluorescence analysis relates to the much publicized mercury contamination of tuna fish. Mercury concentrations at levels of 100 to 200 parts per billion were detected, and, in addition, the ability of this technique to detect other elements at the same time revealed trace amounts of other potentially dangerous contaminants such as arsenic and tellurium. X-ray fluorescence techniques have also been applied to scanning electron microscopes and electron microprobes, which allow determinations of elemental distributions in very small particulate pollutants.

D.5 CONCLUSION

Successful pollution abatement will depend on availability of monitoring instrumentation with readily reproducible precision greater than the permissible pollution limits defined by established standards. Physics research of the past is at the foundation of much of present-day environmental analysis instrumentation. Future physics research can be the source of the considerable improvements needed to support fully the rigid environmental standards that, in the public interest, must ultimately be imposed.

Appendix E: Medical and Surgical Instrumentation

E.1 APPLICATIONS AND THEIR SOCIAL VALUE

In recent years, the medical profession has begun to use modern instrumentation in the treatment and care of patients with results that have often been spectacular.

E.1.1 *Pacemaker*

Probably one of the most widespread applications is the instrument known as a pacemaker, which is implanted in the patient's chest to keep the heart beating in a normal fashion. This is an offshoot of the development of miniaturized electronics and miniaturized batteries that has enabled production of an instrument small enough to be placed in the chest.

E.1.2 *Laser Knife*

The use of a laser, a fairly recent product of physics research, as a "knife" capable of being focused on an extremely small area has greatly improved eye surgery, because it permits the surgeon to cut only where necessary without danger to the rest of the eye.

E.1.3 *Intensive Care Units*

Many hospitals now have intensive care units where modern electronic instruments are used to continuously measure and show the patient's pulse, breathing, electrocardiogram, temperature, etc., so that any change in condition is recognized immediately and can be treated immediately. This has undoubtedly resulted in many lives being saved.

E.1.4 *Mass Spectrometer*

A mass spectrometer, an instrument originally developed by physicists to determine the mass of the atom, is being used in at least one laboratory for rapid identification of drugs in cases of overdoses, thus enabling the proper lifesaving action to be taken.

E.1.5 *Infrared Detectors*

On a somewhat less spectacular note, but still of importance, is the use of infrared detector instruments to provide in a very rapid manner a "temperature map" of a human being, thereby revealing any regions of abnormal temperature that may be indicative of malignancies or circulatory problems.

E.1.6 *Tympanic Temperature Measurement*

A novel and recent application of the well-known Seebeck thermocouple effect has been for the measurement of tympanic membrane temperature during surgery. It has been established that the tympanic membrane responds more rapidly to body temperature change than do the conventionally measured oral or anal areas. Some surgical patients react adversely to anesthesia, manifested by a rapid change in body temperature. Monitoring the temperature of the tympanic membrane thus provides early warning of a developing difficulty, enabling the anesthetist and surgeon to take timely corrective action. The new device is designed to permit safe entry into the auditory canal, and cost has been kept low enough to permit disposal after each use.

E.1.7 *Automated Multiphasic Testing*

A major thrust in general health care is the use of a wide variety of chemical tests as a measure of a person's health. These are forming not only a basis of diagnosis but also a vital part of preventive medicine. The real significance of these tests is that they can be run by the thousands because of the development of modern automatic analytical instruments, thus making the benefits available to many more people. These automatic instruments are often coupled to computers, which can assist in the diagnosis process.

E.2 ACOUSTIC INSTRUMENTS–AUDIOMETRY

E.2.1 *Scope*

The field of audiometry involves instrumentation used in the measurement of an individual's hearing. Two classes of audiometric instrumentation will be discussed: (1) audiometers, which record the subjective response of a listener, and (2) acoustic bridges, used for measurement of middle-ear characteristics.

Recent emphasis on the potential permanent damage by environmental noise to hearing has given rise to a class of monitoring audiometers, which, being almost entirely automatic, provide a rapid, convenient means of screening individuals having potential hearing disorders from those having normal hearing.

E.2.2 *Dependence on R&D in Basic Physics*

Electronic instrumentation by its very nature is a direct consequence of the current technological art, itself a function of research and development carried out by physicists and engineers. The development of the transistor, for example, has made electronic instrumentation, in general, quieter, more reliable, and easier to package. In audiometers, this same technology has resulted in the development of reliable microphones that serve as transducers in speech testing and provide for the calibration of earphones.

Other developments rooted in basic physics that have contributed directly to the design of audiometric instruments include the following:

1. *Standards* Without the ability to make the precise set of acoustic and electronic measurements required for interinstrument reliability, audiometers would probably never have achieved the widespread use they have today.
2. *Transducers* Diagnostic devices that transmit sound stimuli by air conduction and bone conduction.
3. *Feedback* The basic concept of a closed-system feedback loop that has been utilized by physicists and engineers in circuit and system design was borrowed by the instrumentation specialist and adapted to the development of automatic audiometers, which provide much more precise and reliable estimate of hearing than is possible through alternate techniques.

E.2.3 *Social Implications*

The utility of audiometric instrumentation lies in its ability to ensure that the individual will be capable of communicating and, therefore, capable of social interaction. Audiometers and bridges find widespread use in (1) detecting hearing problems before the individual is socially incapacitated, (2) providing accurate diagnoses of the particular hearing deficiency, and (3) prescribing corrective methods to alleviate the disorder. All three uses help to ensure that the individual will be able to function in society.

Appendix F: Nuclear Magnetic Resonance Instrumentation

F.1 INTRODUCTION

Nuclear magnetic resonance (NMR) provides a good illustration of sophisticated basic research that resulted in the development of instrumentation that, through its use in both further research and field measurement applications, has been of major assistance in areas of substantial social benefit. It has benefited the U.S. balance of trade by providing exports not only of NMR instrumentation but also of new products particularly in the areas related to chemistry. The invention, the development of the technique, and its commercialization have been a cooperative endeavor among the pure scientists—in many cases physicists—the applications scientists, the business administrators, and the production workers.

F.2 HISTORICAL BACKGROUND

The phenomenon of NMR was discovered by F. Bloch at Stanford and E. M. Purcell at Harvard in the winter of 1945–1946, which earned for them the joint award of the 1952 Nobel Prize for Physics. The interest of these physicists was the determination of magnetic moments and angular momenta of the various nuclear isotopes. The values of these parameters are important in understanding the structure of the nucleus. The measurement is performed by placing the atoms of interest in an intense magnetic field, exciting them with radio-frequency radiation, and detecting the absorption of energy as a function of its frequency. By measuring the frequency at which energy absorption takes place and relating it to the strength of the magnetic field, it becomes possible to determine the ratio of magnetic moment to angular momentum of the particular nuclear species in question. Since this information provides a characteristic signature not only for the nucleus involved but also, as was soon discovered, for the particular atomic or molecular environment in which it finds itself, the basis for a highly useful analytical tool was at hand.

It was soon found, too, that the measurement could be used in reverse to determine the strength of a magnetic field by relating it to the absorption frequency of a nucleus for which the gyromagnetic ratio was already known. Thus, as a significant by-product, an accurate magnetometer was born.

F.3 APPLICATIONS

The greatest impact of the NMR technique was found in chemistry. It came about because of a discovery, again by a group of physicists, that the resonance conditions for a given atom depended to a small extent on the chemical molecule in which the atom was located. Should several like atoms exist in nonequivalent positions in the same molecule, several slightly different resonance frequencies are found. This finding was designated the chemical shift, because its value is determined by the chemical environment of the atom. Studies of proton NMR spectra have been used extensively since 1952–1953 to determine the structure of molecules and the composition of mixtures and to understand equilibrium reactions. As instrumentation was improved by engineers and scientists, the technique was used for a wide variety of elements, for example, fluorine, phosphorus, boron, and nitrogen, in addition to the original hydrogen, and is now one of the most valuable instruments for chemical research. In fact, after a chemist synthesizes a new product, he can determine within hours many of its structural characteristics by using NMR. In the latest development with superconducting magnets, Fourier transform techniques, and signal averaging devices, the method is being used in the biochemical field to study the structure of very complex proteins for the identification of various hydrogenous and other components of the complex molecule. The present state of development indicates that work with ^{13}C will make it possible to identify the different carbon atom complexes and locations in complex organic molecules, again providing basic structural information for the research chemist and biochemist. The technique is now accepted as a standard method of chemical analysis and is used throughout the chemical industry providing a significant saving in time and cost.

Both NMR and its companion electron spin resonance (ESR), discovered at about the same time by groups at Kazan and Oxford, are used to study biomolecules from simple amino acids and sugars through hormones, enzymes, and DNA. They are used to study molecular conformation and structure, as well as metabolic reaction rates and mechanisms. A recent widely publicized application of NMR and ESR is for detection of a metabolic product of heroin that is excreted in the urine of a user, thus providing one of the first rapid reliable screening tests for drug addiction.

Since about 1955, the magnetometers that resulted from the NMR experiment mentioned in Sec. F.2 have been used to make airborne and ground-station measurements of the earth's magnetic field. These measurements are particularly useful to petroleum geologists in helping to identify formations in which oil might be found. Some of the early earth satellites carried proton magnetometers to determine the magnetic-field strength at large distances from the earth and to relate it to the various theories concerning the origin of the earth's magnetic field. Several companies produce laboratory magnetometers for determining the precise field strength of laboratory magnets.

F.4 ECONOMIC ASPECTS

Estimates for the time needed for the cost savings to pay for an NMR instrument range from a few months to a few years. The usefulness to industrial laboratories has been further demonstrated by the fact that practically all major chemical companies have multiple NMR instruments, some of them being multiple copies of a given type and others having differing characteristics.

The NMR instrument market for chemistry has grown rather steadily since the first sales in 1953. The market size in 1971 is estimated to be $25 million, with approximately 40 percent of the market within the United States. Prior to 1966, almost the entire market was supplied by one U.S. company. At the present time, major suppliers of this market are found in Germany, England, and Japan, as well as in the United States. The ratio of U.S. exports to imports of NMR equipment has dropped from approximately 10 in 1966 to about 1.5 at the present time.

Appendix G: Electron Optics

G.1 INTRODUCTION

This appendix is divided into three parts: historical background, the effect of this type of instrumentation on social welfare, and economic effects. Four instruments have been selected for discussion: the transmission electron microscope, the scanning electron microscope, the microprobe analyzer, and energy-dispersive x-ray analysis.

G.2 HISTORICAL BACKGROUND

G.2.1 *Transmission Electron Microscopes*

It was known in the nineteenth century that the resolution of light microscopes was limited by the wavelength of light. In 1924, as a part of the revolution in physics and understanding of matter stimulated by quantum theory, de Broglie in France proposed theoretically that all matter possessed wave characteristics and that for high-energy electrons this wavelength would be extremely short. In 1926, Hans Busch in Germany made the theoretical discovery that axial symmetric magnetic and electric fields acted as lenses for electrons. Busch verified his theory for magnetic lenses experimentally in 1927 and founded the science of electron optics. In 1932, in the United States, C. J. Davisson and C. J. Calbick experimentally verified Busch's theory for electrostatic lenses.

A period of development then occurred, largely in Germany, where electron microscopes were developed to make images of electron-emitting sources such as heated cathodes. At this stage, the instruments were used mostly by physicists.

The idea of using an electron microscope for examining specimens by transmitted electrons was suggested in a 1931 German patent application by K. Ruderman. Work to develop such an instrument was started by M. Knoll and W. Ruska in 1932, who believed that an electron microscope could, in principle, greatly exceed the resolving power of a light microscope, because of the electron's short wavelength. The first electron microscope to surpass the resolution of a light microscope on specimens viewed by transmitted electrons was built by Ruska in 1934; it surpassed the resolution of the light microscope in 1935. An improved version was later designed by Ruska for the German firm Siemens and Halske and introduced commercially in 1938.

At the same time, independently, A. Prebus and J. Hillier at the University of Toronto designed an improved microscope that the Radio Corporation of America introduced commercially in 1940. It could resolve to 24 Å. In the two decades that followed, RCA sold an estimated 1500 to 2000 instruments in the United States.

Interestingly, electron microscope use and sales have been directly related to the ability of the scientist to prepare specimens that could be studied in this instrument. Specimens have to be thin enough for electrons to pass through and to be imaged by a suitable lens system. The first instruments were largely used to look at plastic replicas of metal and ceramic surfaces and were not widely used to look at biological specimens. This extremely important application awaited the development in the late 1940's of the ultramicrotome by K. Porter at the University of Colorado. This device could cut tissue specimens to the necessary 100 to 200 Å thickness that permitted electrons to penetrate completely. Preparation techniques for biological specimens were developed, and in the 1950's the use of the electron microscope in biological and medical research became widespread.

At about the same time (1948), R. D. Heidenreich showed that it was possible for electrons to penetrate much thicker (1000–3000 Å) metal sections than had been previously thought possible. This led to the direct viewing of metal specimens in the transmission electron microscope and to greater understanding of the morphology and deformation characteristics of metals and alloys. Thus the electron microscope made very essential contributions to the early development of materials science.

Today, transmission electron microscopes operate at potentials up to 1.5 MV and are available commercially with guaranteed resolutions of 2 Å.

G.2.2 *Scanning Electron Microscope*

The basis for a scanning electron microscope was first proposed by M. Knoll, R. Theile, and M. v. Ardenne in the late 1930's. V. Zworykin, J. Hillier, and R. Snyder developed the concept for the forerunner of the modern-day instrument in 1942. However, in neither case was a commercial instrument forthcoming, because the state of technology was not far enough advanced to permit attainment of a resolution high enough to make the technique attractive. It was not until the early 1950's that there was sufficient improvement in electronics and detectors to allow the necessary breakthrough in instrumentation for a scanning microscope to become competitive with other techniques. This was accomplished almost simultaneously by a group at Cambridge University in England and one at Westinghouse Electric Company in Pittsburgh. Early scanning microscopes formed images using primarily backscattered electrons. In the early 1960's, a means for using secondary electrons for imaging was developed that allowed the manufacture of instruments with 200 Å resolution. In 1965, the first commercial instruments were put on the market by the Cambridge Instrument Company in England and the Westinghouse Electric Company in the United States. Initially, sales were not good, and Westinghouse withdrew from marketing scanning microscopes.

However, with the advances made in the scanning electron microscope, it became a convenient instrument for viewing a wide variety of materials at resolutions better than 100 Å and with a depth of focus of the order of centimeters. This depth of focus, coupled with the ability of bending the trajectories of the low-energy (<50 eV) electrons, allows the development of a three-dimensional-like image that has proven useful in the study of the structure of a wide variety of materials (metals, semiconductors, insulators, minerals, plastics, biological specimens, and archaeological artifacts). The ability of the scanning microscope to utilize all types of radiation arising from the interaction of an electron beam with a material (secondary electrons, backscattered or diffracted electrons, absorbed electrons, transmitted electrons, Auger electrons, x-ray photons, light photons, etc.) has greatly increased the usefulness of this technique.

Recently, H. V. Crewe at the University of Chicago has developed a transmission scanning electron microscope using a field-emission electron source that has a resolution of the order of 5 Å. Use of an energy analysis technique on the transmitted electrons has allowed the imaging of single atoms of heavy metals in organic molecules. This system has a great potential and should increase the utility of the instrument markedly.

G.2.3 *Microprobe*

The initial idea of using a fine beam of electrons for the chemical analysis of small volumes of material, using the characteristic x rays generated by the interaction of the electron beam and the material, came from the same Zworykin and Hillier team that developed the electron microscope in the early 1940's. This is another case in which the technology was not sufficiently advanced to allow development of a useful instrument at that time. The developer of this instrument is taken to be R. Castaing, who built one for his doctoral thesis. His instrument was developed into a commercial venture by Cameca in France, and the first such instrument was delivered in the United States in 1958. At the same time, a group* in California, under M. F. Hasler, developed the first U.S. microprobe and went on to become the commercial leader in this field at that time.

G.2.4 *Energy-Dispersive X-Ray Analysis*

Energy-dispersive x-ray analysis commenced with the development of a good multichannel analyzer in the mid-1950's. However, it was not until the development and the application of the solid-state detector by F. Goulding and his associates in the early 1960's that this technique began to look exciting. The detectors developed were the Si(Li) and Ge(Li). At the same time, the development of better and cheaper multichannel analyzers, made possible by technological advances in the semiconductor industry, aided in reducing costs for such instrumentation. Detector technology also advanced rapidly in the late 1960's. In 1965 the best resolution of a Si(Li) detector was 600 eV, while today resolutions under 200 eV are commonplace, and Goulding has reported an experimental system using a windowless detector that has a resolution of 50 eV.

* Applied Research Corporation, later to become a part of Bausch & Lomb.

G.3 EFFECT ON SOCIAL WELFARE

These instruments have had a great effect on the social welfare of the nation. The transmission microscope has been a powerful tool in biological and medical research, especially in the areas of histology and cytology. Today, there is probably no major hospital in the United States that does not have an electron microscope in its pathology department. Transmission electron microscopy has also played an important part in virus and bacteriological research. The conquest of poliomyelitis relied critically on the discovery, made with the help of an electron microscope, that it was a virus disease. Studies on biological macromolecules such as DNA and RNA have been carried out with the transmission electron microscope. These instruments have also been useful in the area of forensic science, and a library of electron micrographs has been established by the Federal Bureau of Investigation for substance identification. The electron microscope has been extremely useful in morphological studies in metals and alloys and has led to the development of better materials. The study of polymers has led to better plastics and fibers. Finally, the instrument is being used to a great extent for the study of particulate matter (<500 Å in diameter) in both air and water pollution.

The scanning microscope has been extremely useful in the area of fracture analysis and has given great insight into fatigue failure problems. Much of the decrease in price of integrated circuits has been made possible by the increased yield of usable integrated circuits obtained from a silicon wafer. This, in turn, can be traced to information obtained from a scanning microscope. For example, in the case of multilevel integrated circuits, it was found that most of the failures (open circuits) occurred in the top level, where the conducting aluminum layer crossed the edge of a lower-level layer. By tapering the lower-level evaporations to eliminate the sharp edges, the yield of these devices was increased more than 300 percent. With the advent of the field-emission gun, the higher resolution thus made available holds promise of revolutionizing biochemical research by permitting the identification of the DNA base sequence in the electron micrographs of genetic material.

The microprobe has been very useful in the chemical analysis of small regions of massive materials. It has been particularly effective in the metallurgical field. For example, the microprobe has provided understanding of the factors essential to high quality plating of an automobile bumper to improve its quality.

The energy-dispersive x-ray analysis system, when used with a radioactive source, represents a portable x-ray fluorescence analysis system. It has been used in this manner in ghetto areas in New York City to determine the lead content of paint on house walls, which poses a threat of lead poisoning. It also provides a portable system for continually monitoring the composition of particulate matter in the air in heavy smog areas.

G.4 ECONOMIC EFFECTS

There are now an estimated 3500 transmission electron microscopes installed in the United States and about 5000 instruments worldwide. The average price paid for these instruments is about $50,000, so they represent a capital investment of roughly $175 million domestically and a $0.25 billion throughout the world.

In spite of the strong early start with a major manufacturer in the United States, American technology did not keep pace with the developments in the field, and now the world market is dominated by Japanese, German, Dutch, and English firms. In 1969, an estimated 335 conventional electron microscopes, representing roughly a $17 million investment, were sold in the United States. The vast majority of these were imported. In 1970 and 1971, an estimated 210 per year were sold. In the next five years, it is estimated that this market will be maintained at about the same level. There has been a spectacularly rapid growth of the use of scanning electron microscopes. Their sales in the United States has gone from near zero in 1969 to an estimated 120 per year in 1971, at an average price of about $60,000. In the next five years, it is expected that this market will more than double.

Initially, the world market for scanning electron microscopes was dominated by its originator, Cambridge. The U.S. market is now split between English and Japanese firms, with effective domestic manufacturers just beginning to develop. The combined sales of the five U.S. companies supplying scanning electron microscopes in 1971 supplied less than 20 percent of the U.S. demand.

There are from 300 to 400 microprobes in the United States at present and approximately 100 more in the world market. The average price is $100,000 per instrument. This market is dominated by two U.S. firms. However, this has been a decreasing market with sales expected to be less than a dozen instruments in 1971. In the next five years, sales should remain about the same as in 1971. It is conceivable that this market may totally vanish if the scanning microscope, used with an energy-dispersive x-ray analysis system, turns out to be as versatile as the microprobe.

The energy-dispersive x-ray analysis system reached the commercial market about five years ago, but sales did not become appreciable until 1969. Since that time, it is estimated that about 300 units have been sold.

A unit would consist of a detector portion with an average selling price of \$6000 and an analyzer portion with an average selling price of \$10,000. Both the detector and analyzer areas are dominated by U.S. firms. It is estimated that in the next five years, this will be a rapidly expanding market.

In the area of balance of trade the electron-optics field when taken as a whole must regrettably be considered negative, even though the microprobe and energy-dispersive x-ray areas have a favorable effect, and even though with both the transmission electron microscope and the scanning electron microscope U.S. companies were among the first to enter the field.

Appendix H: Time and Frequency Measurements

H.1 DIRECT BENEFITS TO SOCIETY AND TO THE SOLUTION OF SOCIAL PROBLEMS

H.1.1 *Improved Navigation*

Time measurement has played an important role in navigation throughout history. For example, celestial navigation did not become possible until an accurate clock capable of operation on ships became available. Indeed, the quest for such a clock–which resulted in the Harrison chronometer in 1729–is itself a dramatic story. In more modern times, specifically during the past 30 years, more accurate and reliable marine and air navigation systems have become available based on groups of radio stations that emit signals controlled by precision oscillators. Initially, these oscillators used very stable quartz crystals (W. G. Cady, 1930). However, more recently atomic oscillators based on rubidium (T. Carver, 1957) and cesium (H. Lyons, 1952) have increased the precision of these systems to higher levels. Navigation systems, such as Loran, Decca, and Omega, use very precisely spaced time intervals, and therefore they depend on clocks to establish the intervals. However, the clocks consist of very precise oscillators, and a means to accumulate the basic time intervals defined by the period of the oscillations is required for the application. A clock, then, is, in this context, simply an oscillator combined with a suitable accumulator.

New navigation systems that depend on radar technology have also been developed, particularly for the navigation of aircraft and space vehicles. These systems are based on the changes in frequency (Doppler shift) of the radar signals caused by the motion of the craft under navigation. These small changes in frequency must be measured with reference to the signals of very precise oscillators in the navigation equipment. Another navigation system, recently developed, is the transit system based on an orbiting satellite.

H.1.2 *Safer Transportation*

Timekeeping and transportation proved to be highly important in the last century and led, in connection with the expanding railroad networks, to the first centrally controlled time-distribution systems. As the speed of transport increased, so did the needs for precision timekeeping and for coordinating these means of transportation. Continually increasing demands have been placed on the precision of the time and frequency instrumentation involved in radar tracking. In general, most transportation systems are unfeasible without some high level of precision time and frequency instrumentation.

These problems are particularly demanding in the case of aircraft because of expansion of commercial air activity, the increasing speeds involved, the tendency to use fixed routes between heavily populated areas, and heavy air congestion around terminal areas. Reliable communication has been one of the most important problems in the past in connection with air traffic control, an important aspect of which involves aircraft-collision-avoidance systems and pilot warning indicators. The related requirement for a large number of channels has necessitated the use of the very-high-frequency spectrum, made possible by the availability of stable quartz oscillators in the airborne equipment, as well as in the ground-based equipment.

An aircraft-collision-avoidance system has been developed, specified, and evaluated, based on a time–frequency concept in which the participating aircraft each carry a precision clock capable of keeping time in synchronism with all the others, such that errors are within a few ten-millionths of a second. A network of ground stations also equipped with such precise timekeepers serves to define a common system time. Radio signals automatically exchanged between participating aircraft and containing the time information make it possible to determine aircraft separation on the basis of the time it takes for the signals to travel. By including precise frequency comparisons of the signals exchanged, it is also possible to determine the relative velocity of the participating aircraft and, thereby, provide the pilots with the logical information they require in order to avoid collisions.

Looking toward the future, the air traffic-control problem will continue

TABLE H.1 Progress in Frequency Standards and Timekeeping Devices

Kind of Clock or Frequency Standard	Date of First Instrumentation	Date of Usefulness	Current Accuracy	Historical and/or Instrumentation Range of Precision	Approx. Cost	Approx. No. in Use	Principal Areas of Application
Sundial	1500 B.C. Egypt 100 B.C. Classical World	To present	- -				
Clepsydra (water clock)	Before 159 B.C. Classical World	To about 1300					Institutional timekeeping
Sand glass	~1300 Europe	~1500					Navigation, internal timing
Balance wheel and foliot bar	~1300 Europe 1729, Harrison	To present	$\sim 10^{-3}$	10^{-2} to 10^{-6}	\$10 to \$200	10^9	Medieval tower clocks, wrist watches, wall clocks, ship chronometers (navigation)
Pendulum	A.D. 1581, Galileo A.D. 1656, C. Huygens	To present	10^{-6}	10^{-3} to 10^{-8}	\$50 to \$2000	10^7	Home clocks and some public time distribution; observatories prior to quartz
Tuning Fork	~1711, J. Shore	To present	10^{-4} (?)	10^{-4} to 10^{-7}	\$200	10^6	Wrist watches

Quartz	~1930, W. G. Cady	To present	10^{-5}(?)	10^{-6} to 10^{-12}	\$400 to \$2500	10^{9}	Communication receivers, secondary frequency standards, portable clocks, transmitter control
Ammonia	1949, H. Lyons	1949 to 1965	5×10^{-11}	10^{-9} to 10^{-12}	Not available	<10	Spectroscopy, low-noise amplifiers (first atomic clock)
Cesium	1952, H. Lyons	1958 to present	5×10^{-13}	10^{-9} to 10^{-14}	\$15,000	10^{3}	Primary frequency standard, timing centers, unit of time, portable clocks (best accuracy), tracking stations, transmitter control, navigation
Rubidium	1957, T. Carver	1963 to present	10^{-9}	10^{-11} to 3×10^{-13}	\$8000	10^{3}	Secondary frequency standard, timing centers, portable clocks, tracking stations, transmitter control, radio astronomy, navigation
Hydrogen	1960, N. Ramsey	1965 to present	10^{-12}	10^{-12} to 7×10^{-15}	\$60,000	~25	Frequency transfer standard, timing centers, tracking stations, radio astronomy, spectroscopy
Methane	1968, Barger and Hall	To present	10^{-11}(?)	10^{-12} to 7×10^{-16}	Not yet available	~50	Spectroscopy, infrared frequency standard, potential transfer standard for length, time, and frequency

to make heavy demands on high-speed communication channels. While voice communications constitute an important part of the systems now in use, it is recognized that communication efficiency can be greatly increased by the application of the most modern concepts of digital communication, which in turn depend on having advanced time and frequency control at both ends of the communication channel.

In addition to communications and collision avoidance, the air traffic-control system also depends heavily on ground-based radar and airborne beacon transponders. Both of these system components also require precise control of frequency.

H.1.3 *Better Communications*

In early radio communication, the efficient use of the available radio frequencies was not a particular problem. As the number of radio transmissions increased, it became necessary to allocate frequencies and define channels. The closeness of the spacing of these channels depends on the precision to which the transmission frequencies can be controlled. The capability to control frequencies was advanced enormously by the development of quartz crystal oscillators.

Radio and television broadcasting as we know it today would not be possible without high stability control of frequency through the use of quartz oscillators. A particularly interesting application of even more precise frequency control has occurred in the case of color television broadcasts by the major networks. Rubidium atomic oscillators are used to control the color subcarrier frequency used within a network in order to avoid undesirable distortions of color during program switching.

A specialized problem is encountered when it is desired to engage in secure communications. In this case, the information is coded by randomly mixing it in precisely defined time and frequency slots. The recovery of this information depends on a knowledge of the coding method used and the availability of equivalent time and frequency control at the receiving end of the channel.

The rapid increase in the use of large-scale digital data communication between computer terminals will in the future require the exploitation of frequency multiplexing and time multiplexing techniques to the maximum extent. Again, the most precise control of the frequencies and time intervals involved will be essential in this connection.

H.1.4 *Better Regulation of Electric Power Distribution*

It is widely known that the regulation of electric power frequencies has made possible the commonly used electric clocks. However, it is not so well known that precise clocks and precise time broadcasts have contributed to better regulation and control of bulk power transfers within complex interconnected networks for the economical distribution of electric power from widely distributed generators. Coordinated regulation of time and frequency is a vital part of the control of power interchange over area interties.

H.2 INDIRECT BENEFITS TO SOCIETY

H.2.1 *Improvement of Man's Cultural Estate*

All the factors already discussed have contributed to the improvement of man's cultural estate. Communications and all of its manifestations including radio, television, stereo broadcasts, telephones, and picturephones have increased man's access to information from widely distributed sources. Access to music, drama, news, and special events has become available to an increasing number of people throughout the world. Looking toward the future, it seems quite clear that such wide access will become even more widespread. The retrieval of information now available in books largely stored in libraries will become possible through digital communications developments. Likewise, improvements in educational systems based on computers with remote terminals including all the audio-visual advantages will also contribute.

Space exploration during the past ten years, in all of its aspects, has also made an impact on man's cultural estate and relies heavily on time and frequency technology. It is also worth noting that the practice of cartography has been improved through the availability of better surveying and distance-measuring techniques growing out of time and frequency developments.

H.2.2 *Scientific Research*

Although it is more the intention of this report to show the influence of physics on instrumentation, a brief reference to the role of time and frequency instrumentation in recent physics research appears desirable. Areas where such instrumentation is utilized include spectroscopy and quantum electrodynamics, lunar ranging, artificial satellites, long-baseline interferometry, pulsars, and planetary satellites.

Many of these instruments and techniques allow a very precise determination of earth angular position as well as of polar motion. The possibility exists of relating these data to geophysical events; in particular, there has been some indication that a correlation exists between the motion of the pole and significant earthquake events. Perhaps such research can provide the possibility of earthquake prediction, with an obvious impact on our society.

H.3 CONTRIBUTIONS TO INDUSTRIAL EFFICIENCY AND PRODUCTIVITY

H.3.1 *Better Instruments and Systems*

In general, advances in productivity require corresponding advances in capability for measurement and control. Some examples of such instruments include a thin-film thickness monitor based on quartz oscillators, the recently developed quartz thermometer, counters and synthesizers in automatic computer-controlled test systems, small-distance interferometry combined with electronic counters in numerical control systems for machine tools, surveying systems based on light beams modulated at controlled frequencies, and telemetry systems for the collection and transmission of large amounts of data in which coding depends on the precise control of time or frequency.

Another area of instrumentation involves radio-frequency spectroscopy. Instruments based on frequency effects in chemical materials (such as nuclear magnetic resonance spectrometers and electron paramagnetic resonance spectrometers, which are more fully discussed in Appendix F) are now widely used throughout the world. These instruments, developed and largely manufactured in this country, depend on the use of stable quartz oscillators.

Precision ranging, surveying, and interferometric measurements of very small distances also constitute areas of instrumentation that depend on the precise control of frequency and time. In the case of the interferometric measurement of distances, the frequency-stabilized laser oscillator is essential.

There are numerous other areas where stabilized frequencies are the basis of important products involving unique and advanced technology. For example, the very high level of short-term stability that is realized in transferred electron oscillators (Gunn oscillators) has recently provided the basis for automatic braking systems on vehicles such as automobiles. In this case, semiconducting materials play an essential role. Such principles will be used more broadly in the future for the detection of motion in general.

H.3.2 *Faster Computers and Teleprocessing*

There is an intimate connection between time synchronization problems, the construction of high-speed computers, and teleprocessing. Measured in terms of bits per second (bps), voice communication between humans can be compressed to 100 bps or lower. The fastest communications channels in operation can run somewhat in excess of 10^6 bps; while based on foreseeable limitations on signal-to-noise levels, frequency stability, and the speed of response of electrons in solid-state materials, 10^{11} bps might appear to be approaching a theoretical upper limit.

The future integration of large numbers of computers into teleprocessing networks is already under way, at least on an experimental basis and seems destined for rapid commercial expansion. There will be an important role of time synchronization in all of this.

H.3.3 *Increased Standard of Living*

Popular sentiment to the contrary, increases in the standard of living in the United States have been brought about to a considerable degree by the combination of the application of scientific and engineering knowledge and a private enterprise business system. Leaving aside the popular contentions of the day, insofar as we believe in some degree of orderly technological advance, the United States can be proud of many inventions and developments in the area being discussed, besides, for example, the availability of television, radio, telephone, picturephone, and private and public point-to-point communication systems (marine, air, and ground). All of these contributions to the standard of living depend on stable oscillators and precision timekeeping in one way or another.

H.4 CONTRIBUTIONS TO INTERNATIONAL BALANCE OF TRADE

In connection with the improvement of the balance of trade, it should be noted that the unique contribution that the United States makes to the world involves leadership in the generation of new and advanced technology. The development in the United States of precision oscillators and clocks constitutes examples of such advanced technology that contribute to the ability of the United States to trade with others. Likewise, the development of advanced instruments such as frequency counters, time-interval measurement instruments, frequency synthesizers, and the systems based on all of these also contribute to the balance of trade. For example, air traffic-control systems, developed and manufactured in the United States, are used widely throughout the world. In fact, the United States has been and still is the world's supplier of the most advanced time and frequency equipment, such as crystal frequency standards, frequency synthesizers and counters, cesium and rubidium atomic frequency standards, and equipment related to receiving and processing standard frequency and time broadcast.

There are also a number of emerging instruments that are closely related to time and frequency technology. These include a very precise thermometer, a precise pressure-measuring instrument, nuclear magnetic resonance spectrometers, electron paramagnetic resonance spectrometers, precision ranging, surveying and interferometric measurements of very small distances, all of which can contribute toward a favorable trade balance.

Looking toward the future, the Josephson frequency-to-voltage conversion effect, at very low temperatures in certain superconducting materials, holds the promise for developing extremely accurate and reproducible voltage-measuring instruments.

Appendix I: Computers

I.1 BACKGROUND

Machines capable of performing mathematical computations have a long history in human affairs. Sophisticated high-speed computing systems, priced economically with respect to the functions they perform, have become available only in recent years. Their availability is based fundamentally on developments in physics in the solid-state domain.

The basic breakthrough constituting the foundation of computer technology was the transistor, invented at Bell Telephone Laboratories by J. Bardeen, W. H. Brattain, and W. Shockley and for which they were awarded the Nobel Prize for Physics in 1956. Subsequent contributions by other physicists relate to magnetic films, ferrites, phosphors, photoelectric devices, and insulators.

Semiconductor technology has evolved from discrete units to integrated circuits and more recently to medium- and large-scale integration, where the components are combined by the thousands on a single small chip, perhaps one eighth of an inch square, to form complete functional assemblies.

Paralleling the development of the computer itself has been progress with the requisite peripheral equipment providing complete overall systems of considerable capability.

I.2 APPLICATIONS

As is well known, computers have found widespread use in business, in engineering, in research, in education, in testing, and in industrial control. Applications in business and engineering are well known, have been widely publicized, and will not be discussed in this Appendix, which will cite a few typical examples in other areas, particularly those that involve other facets of instrumentation.

I.2.1 *In Research*

Analysis of data from an array of gas chromatographs and other analytical instruments is one example of a growing use of computers in research in chemistry and related fields. Typically, the voluminous output information from gas chromatographs is fed into a computer, which performs mathematical calculations and analyses such that the final report is presented in readily understandable, qualitative, and quantitative form.

I.2.2 *In Testing*

Computers are finding widespread use in process control and product testing. An example is the use of computers for sophisticated testing and adjusting of tantalum thin-film electrical resistors, a product that provides stability, precision, and size reduction for modern complex electronic systems. The computerized system determines the value of the resistance and then specifies the additional amount of film that should be deposited to meet specifications of ±0.01 percent. It is reported that many such systems are currently in operation, and it is estimated that manpower savings as a result of their use is between 50 and 75 percent.

I.2.3 *In Traffic Control*

A number of municipalities sense and control automobile traffic in congested areas with the aid of computer-based systems. Such systems adjust the rhythms of traffic signals according to weather conditions and traffic density fluctuations, with programmed interruptions for emergency traffic such as fire-fighting equipment. It is reported that such systems permit substantial increases in orderly traffic flow through business districts.

I.2.4 *In Education*

Some 60 or 70 comprehensive computer systems are reported to be in use for teaching functions in educational institutions. One example involves 29 school districts in 7 counties in Minnesota. Primary grade students and junior high school students use these systems to perform simple arithmetic operations and even to write their own programs. It is reported that one system is used by more than 11,000 students from the fourth through the twelfth grades. In another example of educational use, teachers report that second-grade students using computers become more enthusiastic about subjects such as mathematics as compared with those not using computers.

I.2.5 *In Process Control*

Computer systems for automating industrial processes are being widely applied to a variety of industries such as petroleum, chemicals, pulp and paper, cement, glass, steel, and food. Applications in steel will be the example used here.

Increasing international competition in the steel industry has placed a premium on economic and efficient ore reduction and steel fabrication. Computers are contributing to such process improvement in virtually all steps of an integrated steel mill from the blast furnace to the finishing mills. At each of these steps, extensive use is made of a variety of sensors and transducers, many of which have their origins in basic physics research. The outputs of these units are processed in a computer to develop direct and inferred information concerning the status of the process. The computer also initiates actions that will act to automatically correct deviations from desired levels or conditions in the process. (See Appendix C for further details on steel-mill processing.)

I.2.6 *In Energy Control*

Electric-power interconnection in the United States has developed to the point where a coast-to-coast network comprising 94 percent of the nation's power capability runs synchronously as a single system. At over 100 locations on the interconnection, control assemblies continuously monitor generation and power interchange levels and automatically regulate generation in their respective areas to maintain bulk power transfers, to optimize economy, and to achieve reliable distribution to consumers. At many of these locations, digital computers are being installed, providing a high degree of increased reliability and early warning when system security is in jeopardy. The net effect is improved system operation with minimum threats of brownouts and blackouts and with optimum utilization of available facilities.

I.3 COMPUTERS AND THE BALANCE OF TRADE

There is considerable concern regarding U.S. technology and world trade. Some of the greatest controversy centers around a report written by Michael T. Boretsky, a senior policy analyst in the Department of Commerce, who has specialized in such matters as the technology gap and the relative technological strengths of the United States and the Soviet Union. The focus of Boretsky's concern is the area of technology-intensive manufactured products, which include computers, and scientific instruments. Boretsky notes that these commodities are included in those that are the "most voluminous" in our export trade and that constitute the only group "that has consistently yielded surpluses that have covered the deficits in trade with other commodity groups as well as the deficits arising from other U.S. financial transactions with foreign countries." He notes that the balance of trade in these products improved until the mid-1960's when it leveled off at approximately $9 billion.

Boretsky is concerned about this leveling off and believes that the overall figures mask a rather disturbing trend. The importation of technology-intensive products has, over the past two decades, grown almost 2.5 times as fast as our exports. "The aggregate dollar value of imports reached 55 per cent that of exports in 1969 and, if the growth rates continue as they have in the past, our traditional surplus in technology-intensive goods will soon begin to decline. Within the category of technology-intensive goods, the United States is still doing well in the most sophisticated products, such as computers," Boretsky reported, "but this upper limit of sophistication is growing tinier and tinier."

As stated earlier, Boretsky's analysis is criticized by many economists. However, there seems to be enough validity in his report(s) to attract considerable attention and as a result be widely quoted.

In summary, it is rather obvious that advanced technology, of which computers is a leading physics-based example, has significantly helped the U.S. balance of trade.

I.4 A NOTE FOR THE FUTURE

Advances in physics-related electronics, from vacuum tubes to transistors and integrated circuits, and innovations in memory techniques and in peripherals have all contributed to U.S. leadership in computers. These advances represented fallout mainly from solid-state physics research during an era when this particular branch of physics enjoyed a high level support in the United States. The computer gap, it is believed, has been narrowed appreciably by extensive programs in Britain, Japan, France, and Germany, and possibly in the Soviet Union and China.

More exotic approaches are under development. A laser directed onto the magnetic material MnBi is capable of writing in and reading out data. Storage densities of 2.3×10^7 bits per square centimeter have been achieved, and the estimated cost is only 1/10,000 cent per bit. Another optical memory makes use of a laser to imprint holograms in a small crystal. A multitude of holograms may be stored in one crystal, and each image is recalled when a laser impinges on the crystal at an appropriate angle. Storage densities of 10^{12} bits per cubic centimeter appear to be feasible.

If the U.S. leadership is to be retained, it is essential that the momentum of solid-state physics be maintained with adequate financial support in the years ahead.

Appendix J: Some Future Needs

J.1 INTRODUCTION

Some of the appendixes that precede this one include brief comments on future requirements in instrumentation, some of which, it is hoped, will result from future basic research. Additional needs for the future are detailed in the two following sections of this appendix, one prepared from the viewpoint of an industrialist the other from the viewpoint of a physicist.

J.2 AN INDUSTRIAL USER'S VIEW

A broad range of new instrumentation is needed now to advance science and to meet some crucial requirements of industry. Some of these requirements are engineering-oriented, such as the need for more rugged instrumentation especially in industry, but important other needs are apparent either because existing devices provide a relatively crude measurement or because sensing devices, transducers, and the like are not available for the measurement tasks at hand. It is in the latter two categories that physics has an important and immediate role.

Some important needs for new instrumentation have been identified and are discussed below. The items discussed here are but a sample of the broad range of needs and are offered to demonstrate the scope and importance of new instrumentation requirements.

New sensing devices are needed to control a wide range of processes, to evaluate product quality, and to monitor discharges to the environment. Noncontact sensors are particularly in demand for measurement of flow, temperature, pressure, composition, radiation, and dimensions for a broad range of process-control applications. For example, the conventional orifice meter approach to flow measurement is accurate to only ±2 percent under the most favorable conditions, which in practical applications never exist for any significant period. Instruments of high reliability and accuracy are sorely needed for these types of measurements. The nondestructive inspection of materials presents a fertile field for new approaches as well.

The pressure for dramatic improvements in product reliability will continue to heighten the need for devices for measurements of composition, dimensions, and surface and internal quality of materials in intermediate stages of manufacture and in the final product. For example, high-speed remote inspection for sheet defects in the metals industry is an important problem that has not been solved. In the pharmaceutical industry, for instance, powder properties such as wettability, lubricity, and flow characteristics cannot be satisfactorily measured although knowledge of these properties is important to the control and improvement of processing. The satisfactory solution of these types of instrumentation problems would yield substantial returns by way of productivity increases in industry. However, in many cases new concepts for sensors must be forthcoming from physics research to provide the basis for instrument development.

While many devices exist and are in use for chemical analysis, there are many problems that cannot be solved by the existing instrumentation. The analysis of certain raw materials flows into and product flows out of processes is an important example. Furthermore, certain *in situ* chemical-analysis measurements of materials in process, particularly in severe environments such as in oxygen steelmaking, would result in large gains in productivity. Pollution monitoring devices for the reliable continuous measurements of low concentrations of elements and compounds discharged to the atmosphere and to streams and lakes are sorely needed. The specific ion electrode has been particularly useful for this latter class of applications but is a solution to only a small part of the total problem.

Important new approaches are needed for instruments in connection with diagnosis and therapy in the health-care field. Great strides are currently being made in this field, based on relatively new physics, but cell-identification presents one important case of the inadequacy of instrumentation. Society has assigned, and will continue to assign, a high priority to preserving human life and reducing human misery. Instrumentation plays a key role in effective responses to this need, but the advances here are closely allied to new concepts in physics.

J.3 AN APPLIED PHYSICIST'S VIEW

Summarized in this section are some recent developments of physics that show promise of exerting important commercial and social impact. It is, of course, too early to evaluate the full potential of this impact or to identify those approaches that will fail to displace existing technologies. Nevertheless, if successful, some of these approaches may exert tremendous leverage.

J.3.1 *Acoustic Holography*

Acoustic holography makes use of acoustic waves to form images of "objects" immersed in optically nontransparent media. Alternatively, acoustic holography allows one to look inside objects, as is possible with x rays. However, it offers both advantages and disadvantages relative to x rays. Although acoustic waves appear to present less hazard to biological tissue than do x rays, acoustic holography lacks the high resolution attainable with x rays and, hence, in some cases cannot "see" in such fine detail.

Several approaches to acoustic holography have been developed. In one variation, two acoustic transducers are driven in phase by the same power source. They are immersed in a liquid and are so oriented that their acoustic beams combine at the surface of the liquid. This surface is distorted in accord with the acoustic-beam-interference conditions. When the object under investigation is inserted between one of the acoustic transducers and the surface, information on the internal character of the specimen is carried by the acoustic beam as it approaches the surface. In this instance, the resulting liquid-surface distortion amounts to a phase hologram. Coherent light from a laser source reflected from this surface forms a visual image of the internal structure of the object.

Preliminary studies indicate that acoustic holography may provide views of vital human organs not obtainable by conventional methods. In industrial applications, acoustic holography comprises a unique means of "seeing" internal voids, cracks, delaminations, and other flaws in structural components.

J.3.2 *Ion-Implantation Doping of Semiconductors*

Until recently, the impurities or dopants that must be introduced into semiconductor materials in order to produce useful transistors, diodes, and microelectronic circuits have been added only by alloying or by diffusion techniques. However, over the past few years instrumentation and techniques have been developed for doping semiconductors by energetic ion bombardment. In this technique, called ion implantation, the dopant or impurity atoms are literally shot into the host semiconductor material. This process has several advantages over conventional doping methods. The precise control achievable with ion-implantation instrumentation allows the desired dopant distribution to be controlled within tighter tolerances. This means, for example, that the metal-oxide-semiconductor field-effect transistor (MOS FET), mainstay of the microelectronics industry, can be made with less leakage capacitance between gate and source or drain and hence can be made to operate at much higher frequencies. Ion-implantation techniques can also produce subsurface doping and other specialized doping profiles that allow fabrication of new types of devices, devices that would be quite impossible to make by diffusion techniques.

The economic production of ion-implanted devices has been heavily dependent on the development of efficient instrumentation, ion sources, and ion accelerators.

Scientific feedback in this field has led to vastly improved techniques for determining the locations of impurity atoms in host lattices. In turn, this knowledge provides a basis for achieving the desired degree of electrical activity for the dopant ions.

J.3.3 *Electron-Beam Technology in Microelectronics*

Current microelectronics processing technology makes use of optical masks for exposure of photoresist to delineate the intricate patterns of microelectronic circuits. The limit on the miniaturization of microelectronic circuits by this method is set by the wavelength of the light used to expose the photoresist. If the pattern of interest has dimensions of the order of the wavelength of light, diffraction effects become appreciable and preclude attainment of a precisely defined pattern. However, electron beams may be focused to dimensions a factor of more than 20 times smaller and hence may be used to expose electron resist patterns on a much finer scale. Instrumentation for this approach to microelectronic circuit fabrication is currently under development and could lead to planar circuit-element densities many times greater than are achieved by optical processing. Combined with ion-implantation doping techniques, this electron-beam approach shows promise of leading to much faster and more compact large-scale integrated circuits for communications systems, computers, and consumer electronics.

J.3.4 *Magnetic-Bubble-Domain Computer Memories and Shift Registers*

Magnetic materials research has led to the development of thin layers of garnets and orthoferrites in which small regions or domains of opposite magnetization may be established and by suitable techniques may be transported stepwise along a desired path. Such domains are called magnetic "bubble" domains. The presence or absence of such a bubble may correspond respectively to the 1 or the zero of a binary logic. Hence magnetic bubbles may be used for the storage and recall of information.

Although still in the development stage, bubble-type computer memories and shift registers have already been fabricated and tested. Particularly exciting is the potential for supplanting existing disk memories. The bubble memory should be able to provide a volume reduction of a factor of about 1000, a power reduction of a factor of about 50, a speed increase of a factor of about 10, and cost reduction of a factor of about 10. Also holding promise is the large-scale application of magnetic-bubble-domain shift registers to telephone dialer systems.

J.3.5 *Microwave Signal Generation and Processing Instrumentation*

The development of purely solid-state means for generating, amplifying, and processing microwave electromagnetic signals has been a significant outgrowth of physics research. This, in turn, has led to the development of sophisticated new microwave instrumentation. Progress in this direction was first evident in the invention of the Gunn oscillator (a transit-time microwave-generating device) and the Gunn amplifier, both of which utilize the compound semiconductor GaAs. These devices are smaller in linear dimensions by many orders of magnitude than earlier microwave devices and are also much less expensive. However, a complete microwave system or instrument often requires in addition waveguides and signal processing components, which may be both large and expensive. Thus, advances were required in these signal-processing components as well if the newly achieved Gunn devices were to be fully exploited. Very recently, research on the physics of surface-wave acoustics and magnetic surface waves has led to the development of miniature pulse-compression filters, tunable filters, programmable tapped delay lines, digital filters, and other microwave signal processing hardware. It will be possible to couple such acoustically based components with Gunn devices to produce integrated microwave circuits. This will represent an advance in microwave instrumentation quite analogous to the step, already achieved, when microelectronic circuitry replaced conventional discrete-component circuitry in the lower-frequency range. The basis for the compactness of the new microwave technology lies in the fact that acoustic microwave devices may have linear dimensions smaller than their earlier electromagnetic waveguide counterparts by the ratio of the speed of sound to the speed of light or about 10^{-5}.

The new miniature microwave technology can yield a multitude of inexpensive consumer-type instruments, e.g., small radars that may be used in automobiles, in a blind man's "cane," and in burglar-detection systems. The new technology is already being incorporated into sophisticated military radar systems.

J.3.6 *Josephson Effect Instrumentation*

This new class of instruments makes use of one or more Josephson effect junctions. Such a junction may consist of two superconductors cooled to temperatures near absolute zero and separated by a very thin (about 10 Å or about 10^{-7} in.) insulator or by a weak superconducting link or point contact. In accord with theoretical predictions by the British physicist Brian Josephson, such a junction may act as a very sensitive current or voltage detector, a very accurate voltage standard, a generator of high-frequency radiation, a detector of millimeter-wavelength radiation, or a very sensitive detector of magnetic field. As developed in a number of laboratories, devices utilizing the Josephson effects are just now finding commercial and military applications.

So-called quantum interferometers consisting of two Josephson junctions in parallel are the most sensitive detectors of magnetic field known. They are so sensitive, in fact, that the magnetic field generated by the human heart is readily detectable. Magnetocardiograms have been rather routinely obtained in magnetically shielded rooms by this technique, and a commercial instrument for producing magnetocardiograms is available for a few thousand dollars. This approach has the advantage that the patient need not be contacted with electrodes, which not only speeds the data-acquisition process but also eliminates the (admittedly remote) possibility of electrocution. To date, the medical evaluation of this remarkable development has not been aggressively pursued, and so it is not known whether magnetocardiography will ultimately supplant or will merely augment conventional electrocardiography. However, it is already clear that screening of large numbers of people may be conducted much more speedily by magnetocardiography.

The very high sensitivity to magnetic fields of the magnetic quantum interferometer will, without doubt, also find wide application in earth-resource prospecting inasmuch as magnetic anomaly mapping has long been used as a means for locating natural resources. Magnetic anomaly detectors have also found widespread use in submarine detection, and thus, magnetic quantum interferometers should add a new dimension to anti-submarine tactics.

Because the dc voltage that may be established across a Josephson junction is related directly in terms of constants of nature to the frequency of the ac Josephson current flowing in the junction, the Josephson junction provides the most precise and reproducible voltage standard in existence. Thus Josephson devices are finding application in national and international primary standards laboratories.

Josephson junctions may be used to detect voltages as small as 10^{-16} V. Thus far, utilization of this extreme sensitivity has been confined primarily to purely scientific investigations.

Finally, the Josephson junction is perhaps the most sensitive detector of far-infrared radiation known. It has already found significant application in astronomical observations. Moreover, because of the many military applications of infrared-radiation detectors, the Josephson junction will doubtless find application in military systems.

XIII
Education

A theoretical science, unaware that those of its constructs considered relevant and momentous are destined eventually to be framed in concepts and words that have a grip on the educated community and become part and parcel of the general world picture—a theoretical science, I say, where this is forgotten, and where the initiated continue musing to each other in terms that are, at best, understood by a small group of close fellow travellers, will necessarily be cut off from the rest of cultural mankind; in the long run it is bound to atrophy and ossify. . . .

ERWIN SCHRÖDINGER (1887–1961)
British Journal for the Philosophy of Science
Volume III, pp. 109–110 (1952)

One is constantly asked, When should this scientific education be commenced? I should say with the dawn of intelligence. As I have already said, a child seeks for information about matters of physical science as soon as it begins to talk. The first teaching it wants is an object-lesson of one sort or another; and as soon as it is fit for systematic instruction of any kind, it is fit for a modicum of science.

THOMAS HENRY HUXLEY (1825–1895)
Science and Education
D. Appleton & Co., New York and London, 1893, p. 128

PANEL MEMBERS

RONALD GEBALLE, University of Washington, *Chairman*
E. LEONARD JOSSEM, The Ohio State University
JOHN H. MANLEY, Los Alamos Scientific Laboratory
EUGEN MERZBACHER, University of North Carolina
PHILIP MORRISON, Massachusetts Institute of Technology
LYMAN G. PARRATT, Cornell University
GEORGE T. REYNOLDS, Princeton University
ROBERT H. ROMER, Amherst College

CONSULTANTS AND CONTRIBUTORS

J. HAMILTON ANDREWS, National Science Foundation
ARNOLD B. ARONS, University of Washington
WILBUR V. JOHNSON, American Association of Physics Teachers
JUSTIN C. LEWIS, National Science Foundation
MORRIS LERNER, Barringer High School, Newark, New Jersey
BERNARD E. SCHRAUTEMEIER, Meremac Community College, St. Louis, Missouri
W. MICHAEL TEMPLETON, University of Washington

LIAISON REPRESENTATIVES*

MARCEL BARDON, National Science Foundation
CHARLES H. CARTER, National Aeronautics and Space Administration
EARLE W. COOK, Atomic Energy Commission
PAUL F. DONOVAN, National Science Foundation

* In general, no agency was represented on the Committee or any of its panels by more than one person at a given time.

Preface

The United States has never developed a rationale for determining what constitutes a harmonious balance between its scientific enterprise and its total level of activity.[1] Prior to the Second World War, such considerations were excluded, almost as a matter of principle, from the concerns of government. By 1939, through the circumstances of its internal development as a science, physics had acquired principles and techniques that could be and were brought to bear on warfare and communications with great effectiveness. It seemed to offer a model of basic science aiding national policy; during the postwar decades government and industry saw great material advantage in nurturing its growth. Soon all of science was enjoying increased support and prestige. However, the notion that socially and economically imposed limitations might follow, although entertained by some, was never articulated forcefully. The scientific community never felt compelled to weigh as a matter of practical importance the implications of its accelerated growth rate. Thus today, with the nation facing social and economic problems of great immediacy, we still lack a national policy toward science.

The Physics Survey, of which this report is a component, has been undertaken at a time when the evident change in public attitude and political stance toward science is being discussed vigorously. The basis for a fruitful relationship between science and public policy remains, however, elusive. The formulation and acceptance of a national science policy is more likely to emerge when the public and scientists alike gain a more rounded understanding of the scientific endeavor. The problem appears to be in essence an educational one; starting at the earliest level of our school system, science must be depicted as at one and the same time an expression of a persistent human desire for knowledge and a tool whose use brings some predictable and many unforeseeable consequences. To help our technologically based society retain, or even regain, its balance, scientists must make a far more thorough commitment to education than ever before and must also make a searching appraisal of the social implications of their own activities.

Technology, emerging from an incubation period of several thousand years, has undergone an apparent metamorphosis; we are now confronted with the necessity of regulating its manners and growth if we wish a habitable, steady-state earth. Guiding ourselves and our technology into this unaccustomed region will be a delicate business requiring a nice balance between stifling restraint and reckless laissez faire. Science, which society has been accustomed to viewing as the source of new technology, will now play a large role in providing tools for managing technology. Energy, matter, time, space, and even entropy will be common concerns in the decades ahead, in our politics and economics, in our national policies and international negotiations. Society, the ultimate manager of technology, must comprehend the several roles of science and, as well, some universally applicable scientific principles.

A truly comprehensive treatment of educational needs falls outside the scope of a survey of physics alone, even outside science as a whole. Nonetheless, the times call for the practitioners of physics to face up to the full range of their educational responsibilities. These responsibilities transcend the education of future scientists.

A faithful survey of the professional activities of physicists should acknowledge the initiative of those who have already taken strong roles in the present wide variety of ventures aimed at improving the quality of education, but it should also assess existing shortcomings and project the needs of the future. It must make the individuals of today's physics faculties uncomfortably aware that the responsibility for improvement is *theirs* and cannot be left to a new generation, for during the next decade or two not very many new faces will join their ranks.

In accepting a commitment to provide a treatment of physics and education that will complement the other reports making up this survey, we have been acutely aware of the complexity of the topic, of the difficulties of adequate comprehension and statements of objectives, and of the vast number of questions that can be raised about the machinery and effectiveness of physics education. We fully realize that the social and economic underpinnings of the entire educational system, especially the higher educational system, are changing as we write and are likely to be transformed drastically during the next decade or so. We have accepted this charge knowing that there are more questions than answers but believing that many factors in the current social scene make such a survey timely and important. Our approach is interpretive and critical, a reflection of the mood of today.* It does not, however, abandon all contact with the present scene. We assume America will not green so uniformly during the next decade that schools, colleges, and universities will grow unrecognizable, that the basic scientific disciplines will blur and die, and that man's motivation to learn will be lost.

We have perceived our task to be the treatment of the most basic and universal role of physics in education and of the education of physicists. A different point of view would have stressed certain contemporary problems such as the dearth of women in physics and of physicists from underprivileged backgrounds and the educational imperatives these imply. This point of view would have led to an extensive treatment of the special needs of different kinds of educational institutions, including, for example, women's colleges, black colleges, two-year colleges, liberal arts colleges, and emerging state colleges. We also could have given much more space to a description of the manpower problem we are experiencing, to the various projections of future needs of educational institutions and industry, and to their quantitative implications for education. We sense that many are speaking to these matters, while there is little opportunity to put forth, in the framework of a comprehensive survey of physics, a broad view of the role of physics throughout the educational system.

We come to the task as physicists, varied in our professional interests and experiences. Against the background of the vast educational scene we are attempting to cover, however, we sense ourselves a homogeneous, limited group. We have attempted to exploit the strengths this brings and to compensate for weaknesses by expanding our report through contributions, comment, and criticism from others—students, teachers of many kinds, other scientists, administrators, and lay individuals. We have been especially aware of limitations as we have treated education in the elementary and secondary schools and when we have extended our concerns to what is more properly science education than physics education. Even within the

* During 1971, a study and Conference on Priorities for Undergraduate Physics Education was conducted by the Commission on College Physics. A number of the issues we discuss were also treated then, in the light of background papers prepared by the Commission staff. We have not attempted to duplicate their effort, which collected and summarized a large amount of reference material and numerical data. A preliminary account of the study and conference appeared in Newsletter No. 25 of the Commission, November 1971. The complete report appears in a special issue of the *American Journal of Physics* in 1972. Newsletter No. 25 also gives an account of a Symposium on the Education of Physicists held at the Battelle Seattle Research Center in 1971.

realm of the efforts made by physicists to innovate and to improve, our knowledge and surely our wisdom are narrowly circumscribed. All in all, it would be incorrect to represent this report as other than our collective commentary on the educational scene as it has developed from our individual past experiences and current discussions. We offer it as a point of departure rather than as a definitive pronouncement.

Acknowledgments

This report owes much to a number of people who have contributed supporting papers, suggestions, and criticism during its formulation. Foremost among them are Arnold B. Arons, Wilbur V. Johnson, and Bernard E. Schrautemeier, who prepared statements basic to Chapters 2, 7, and 4, respectively.

Senta A. Raizen, Justin C. Lewis, and John H. Andrews of the National Science Foundation helpfully supplied data. Robert F. Tinker, then of the Commission on College Physics, and Rufus C. Salyer of the University of Washington College of Education provided additional data. Albert M. Clogston, George Ivany, and Morris R. Lerner wrote helpful statements for us. Letters giving useful information and points of view were written by William A. Blanpied, Thomas Carver, Ronald D. Edge, Herbert I. Fusfeld, Joseph A. Giordmaine, W. D. Hardin, Robert N. Little, Walter E. Massey, John Mays, F. G. Shoemaker, R. H. Socolow, Arnold A. Strassenberg, Robert J. Wilkinson, Robert M. Williamson, and Rolf G. Winter. Nathaniel H. Frank offered his views in an interview. Frank Oppenheimer and W. Michael Templeton participated in a meeting of the Panel, and Mr. Templeton compiled information and read early drafts of portions of the report. Roger G. Olstad suggested sources of information. W. E. Haisley helped with editorial revision. Finally, Essie Levin spent much time supervising and participating in the typing and collating of drafts and keeping track of Panel activity, much of which took place after hours.

We are most grateful for the help these individuals have provided. Since in many cases they will not recognize (or perhaps approve of) the form taken by their contributions or the use made of their advice, we do not ask them to share in the responsibility for the contents or the tone of the report. If we have omitted any contributors, it is by no means intentional.

Contents

XIII
EDUCATION

1 Recommendations

In keeping with the character of this report, our major recommendations to the Physics Survey Committee are broad and qualitative. We have chosen to single out those that we consider to have the most far-reaching importance and present them here. However, the report contains many other suggestions that we believe to be significant for the strengthening and extension of physics education throughout our nation.

1.1 THE SCHOOLS

Science should be taught in the schools beginning with the earliest grades. It should be taught in a manner that encourages inquiry by the child, independent and self-paced, but guided. Several examples of curricula designed in this mode are available at all school levels either for adoption or to serve as models for subsequent development.

Successful use of inquiry-directed instruction requires teachers who have themselves learned to investigate in this manner. At present, the education of teachers is very weak in this respect. A broad and intensive effort is needed to give prospective and in-service teachers the background for leading pupils into independent inquiry.

Science teaching is promoted by classrooms that offer freedom of action and a variety of tools and materials, not necessarily elaborate or expensive, for free use. It also is promoted by flexibility in the use of time, so that varying degrees of intensity can be accommodated.

The advising of students respecting either careers in science or the value to be gained from a course in science should be carried out by counselors who have a better understanding of the nature of science and of the kinds of ability and personality traits that lead to success in the variety of scientific careers. Counselors who possess this understanding will be more likely to assist girls and children from minority or underprivileged backgrounds in developing an interest in science. They also will be better able to identify and encourage scientifically talented children of whatever background. Recent evidence pointing to the early age characterizing an option for science emphasizes the need for improved counseling. Agencies of the physics community should seek ways to help counselors inform themselves about science and scientific careers.

1.2 UNDERGRADUATE EDUCATION

To meet the goal of increased scientific literacy among the general public, physics faculties should give far more attention to courses for the general student. The intent of such courses should be to promote understanding rather than to cover a wide range of material, implying that the pace may be modest, even at the cost of omitting the latest developments in physics.

We advocate widespread introduction of courses conducted in the inquiry mode and intended for elementary and secondary school teachers. These should make extensive use of simple experimental materials and situations, similar to those encountered in the schoolroom. Physics faculty members should seek the cooperation of the education faculty to encourage the population of these courses. They also should acquaint themselves with developments in the psychology of learning.

Students who enroll in physics courses to fulfill requirements of other disciplines are a serious responsibility of physics departments. We recommend periodic consultation with faculty from these other disciplines, in order to make such courses as interesting and valuable as possible.

Students majoring in physics should be encouraged to understand and explore the relationship of their subject to other fields of knowledge. The notion that undergraduate concentration in physics is a useful foundation for other studies and for a variety of careers should be implicit in all physics major curricula.

The development of curricula in technical physics is important to the education of the large number of technicians needed by industrial society. Cooperation between the physics faculties of the two-year and the senior institutions is needed, not only to develop adequate curricula of this kind, but also to ensure articulation of the appropriate segments of their instructional programs.

1.3 GRADUATE EDUCATION

The notion of a core of knowledge that emphasizes and transmits physics as a unified discipline should be maintained. The material making up the core will change slowly with time but should always permit ready entree into topics characterized by having major unsolved problems and should retain a balanced treatment of macroscopic and microscopic phenomena.

Graduate programs should promote the concept that physics is a broadly applicable subject and that the thesis requirement is not intended to impose a long-term commitment on the student. A prolonged period spent in thesis research is antithetical to this concept.

Since opportunities for careers in the application of physics promise to one day outnumber those for careers in teaching or pure research, an intermediate degree appears desirable. To this end, physics departments should give attention to reshaping the master's degree or introducing an alternative one.

The opportunity to serve as apprentice teachers and to gain background for teaching careers should be made available to all graduate students.

Recalling that many outstanding physicists began as students of other disciplines, physics departments should be willing to accept into their graduate programs those who show interest and promise even if they lack aspects of the formal requirements for admission.

1.4 FEDERAL SUPPORT FOR SCIENCE EDUCATION

Federal support promises to remain an essential ingredient of innovation and self-renewal in science education in the near future. We note a progressive decline in this support, as exemplified by the National Science Foundation's education budget, during the past few years. We ask the reversal of this trend and commend those members of the Congress who advocate a strong federal role in science education.

The improvement of science education, through innovation and self-renewal, should be carried on as a component of science itself, thus insuring the involvement of scientists. At the present time, the National Science Foundation is the agency best equipped to manage the federal responsibility for science education, because its staff comprises scientists;

communication between the federal government and the science educational sector is thereby facilitated. We advocate a continued leadership role in science education for the Foundation and commend actions taken by the Congress to maintain and strengthen this role.

2 Introduction

As the chief inheritor of centuries of human inquiry called Natural Philosophy, physics (together with its close relative, astronomy) has had a long time over which to influence the thoughts of humanity. In addition, at an early date it became a science that concentrated on relatively simple systems. Here, in contrast to the systems of chemistry or biology, the number of variables is limited and experimentation yielded quantitative information which first conveyed the far-reaching insight that physical phenomena operate in accordance with mathematical laws. The result was a rapid achievement of description and understanding which laid a firm foundation for an increasing ability to handle more complex phenomena.

The continuing interplay between physics and practical arts has been and remains a significant factor in the development of both. In many ways physics could have lost its influence on thought or on efforts to benefit in practical ways from knowledge of nature. That it has not done so is due in no small part to its continuing concentration on elemental phenomena together with its willingness to apply its learnings and to aid other sciences and engineering in pursuing lines of investigation which were originally a part of the main body of physics. A second reason arises from more recent history. World War II and the subsequent years brought physics and the physicist under public scrutiny. The result has been a resurgent involvement with public affairs, and with society at large. Consequently, the thoughts of the physicist on education are strongly conditioned both by his subject and by the influence of the social environment in which he lives.

Much has been written in recent years about the growing influence of science and technology on the everyday world. Practically any shade of opinion respecting this development can be found, from a characterization of science and technology as one culture and presumably everything else is another, to science as a new religion, or to the physical sciences as pacemakers of civilization. This variation of opinion is traceable to widely different value judgments, but there *is* a consensus that science and technology constitute a still-mushrooming social force. Physics as a science and as a basis for technology is deeply involved. Education, including physics education, must reflect the forces of society and react to changes in that society.

Although as physicists we have an ingrained desire to approach a topic with precise definitions and other paraphernalia of our trade, we have to realize that education may not be amenable to such procedure. We, like others, have to be content with an indirect approach to a definition of content and objectives in a value-laden subject. We begin our survey by making some general comments on education and proceed in a more operational sense, by describing humanistic expectations from the study of physics.

Among all the qualifying words used to describe education—liberal, vocational, professional, technical, for citizenship, for life, and so on—there seems to be a common denominator: We are led to postulate that an underlying purpose of education is to increase the individual's awareness and understanding. This includes awareness of one's own nature as observed from one's own actions and reflections, awareness of one's fellows as expressed by the visual and auditory arts (painting, sculpture, drama, literature, music, language), and awareness of one's interaction with other living things as learned from the social, behavioral, and biological sciences. It means awareness of the inanimate universe, one's place in it, and one's ever-increasing involvement with things derived from knowledge of the materials and forces of nature. All these kinds of awareness provide guides for the conduct of one's own life. They are boundary conditions and at the same time expansive horizons that can define personal freedom. Their influences cover the most practical aspects of one's livelihood and extend to the deepest reaches of one's spirit.

Since the pursuit of physics always has been related to technological development, it could be assumed that the main purpose of physics education is to further our control over nature for our own immediate benefit. Such a view is both narrow and false. It ignores the most significant fact, that it is the awareness of humanity of which we speak. The subject, physics, is man-made by individuals who devote themselves to understanding the physical universe. Far from complete, it retains the quality of a quest. This was understood clearly by Martin Green, a literary critic and teacher who, challenged by the "Two Cultures" debate of a few years ago, wrote after completing courses in mathematics, physics, and biochemistry[2]:

> We need a new humanism, to show how a man can know what Rutherford knew and what Eliot knew, and can become, not despite but because of that knowledge, fully a citizen of our world. We need to break the tabu of incommunicability that has been laid on that knowledge; to learn to transmit it to those with equal though different

intellectual experience. The belief that a non-scientist cannot achieve any significant understanding of science must be dismissed as a delusion, a symptom of the disease itself. The difficulties are of course great. But they cannot be insoluble. For the problem is not to acquire a certain amount of information, or even understanding, but to employ a certain amount of serious attention. It is *our* activity, the way we act on our tiny section of physics or chemistry, the depth of the science we know, that is important. What we need is intellectual engagement, imaginative participation at the same level of intellect and imagination as we know in our own subject; though nine-tenths of this material must be taken on trust. A good teacher of science can do that in a year starting literally from scratch, just as a good teacher of literature can.

Although Green was writing for mature students of literature, and by the "level of intellect and imagination as we know in our own subject," he meant the university level, his comment is no less pertinent for all levels. To teach physics in the manner he asks poses a fundamental difficulty at any level. Physics is inherently a quantitative discipline and an experimental one. It always has been a costly study, as was recognized two centuries ago by Joseph Priestley when he wrote: "Natural Philosophy is a science which more especially requires the aid of wealth."[3] At any given time, the concepts employed by physicists at their working level are abstract and removed from common experience. Until the present time, educational practices in this country have been prompt and effective in sorting out (or creating?) two populations, roughly speaking, one of which has a bent toward quantitative discipline and another larger one that tends to shun the quantitative and instrumented. We shall subsequently encounter examples of initiatives that promise to circumvent and overcome these difficulties but that require for widespread effectiveness the active participation of college and university faculties.

Physics, because it integrates the philosophical, the quantitative, the ingenious, and the practical, has potential appeal to many facets of human interest. It is in this sense that physics is a humanistic subject, not only for its practitioners but for all men to the extent that their own concepts, their awarenesses, even their fancies are fashioned by the world view given by science. We shall return to this idea later, for it is a key to the way in which physics education responds to the societal forces of modern science and technology.

To continue our operational description of some of the purposes of physics education, it will be convenient to define more explicitly the two populations mentioned above, although one must be careful to avoid dogmatism. In one category will be the eventual researchers, science teachers, scientists, and engineers; in short, all individuals who will participate in some way in meeting society's need for pursuers and manipulators of science and technology. For all such, but with a wide variation of intensiveness and content, physics education will emphaisze the skills and the subject matter of physics but must not obscure the humanistic aspects. In the other category will be those individuals who do not contemplate careers in science and technology; we include all who seek to learn at any age, whether through formal or informal means, and those who gain information about the physical world merely through casual encounters. For them, too, we believe there is an appropriate study of physics. What is important for them and why is considered in the next chapters; here we merely observe that to achieve its purposes, the content and approach of physics education must surely be different for those two broad groups.

Physicists have always been somewhat uneasy about their efforts to bring their subject to the public, but the forces of today threaten to turn this malaise into an acute illness. The physicist sees a growing alienation from commitment to rational thought, an alienation that he believes can only be harmful to society in the long run. He believes the corrective to be effective science education for the public. Should and can the physicist participate in this immense task? Can he afford to neglect it? Should he wish to participate, how and where can he make himself effective? With notable but sparse exceptions, physicists have not proved to be enormously successful in ventures of this kind because their motivations, and their education as well, have had other primary objectives. The feeling of malaise originates, too, from uneasiness about the nature of education for professional scientists. Has the physicist educated yesterday been attempting to perpetuate his own image without appreciating the changing attitudes of the present and the far less certain demands of the future? Is it truly the obligation of the scientist to teach something beyond science, namely, concern for the future and concern for the unpredictable uses of science? Even granting that it is his obligation, how can it be done? Prompted by the warning given by Schrödinger that is quoted at the beginning of our report, we include reflections on these matters among the considerations of the report.

Whatever we may conclude about the purposes of physics education, its approach and content should take into account its relationship to the entire educational endeavor, itself integrated into the whole social structure. We are attempting to deal with a small part of an extremely large, nonlinear, and intricate dynamical system. Its substructures and components affect one another in a variety of ways and at differing rates. A change effected at one point will inevitably manifest itself later elsewhere in the system. Limitations of time and space and, above all, of our own understanding of the networks of interaction among the parts of the system oblige us to try to isolate some subsystems so we can attempt to treat them individually. We suppress the complexity at our peril, proceeding as best we can to understand and improve the educational structure we have created and on which we rely so much.

3 Physics in the Schools

"Eppur si muove," muttered Galileo as he signed the great recantation. Or so the legend goes. It is a compelling legend, and every physicist feels some pride and some sympathy in the telling of it. For it is the universe, he thinks, that decides about these matters, not man; yet at the same time he recognizes the power of societal compulsion. There is no better allegory for the inward nature of our science today than that one. Men are not free of the influences of the sixteenth century yet.

How far has the model of the world that physics triumphantly describes penetrated into the "subsoil of the mind"? It is easy to show that most educated adults do not in fact share Galileo's confidence. They know by perception that the sun moves in the sky; they "know," too, by force of the climate of scientific opinion, that it is really the earth that moves. But left to themselves, it is the first model they might place their bets on. Perhaps the moon voyages are beginning to permeate even the subsoil, but one cannot be sure.

The task for physics, if it is to survive in health and pride of accomplishment, is to achieve in the nonscientific community a genuine understanding, in place of the straightforward preconceptions of Aristotle, of the way in which scientific investigators have wrested other and deeper pictures from nature. That understanding cannot in fact be found outside the narrow layer of the population who practice either academic science or the technical crafts, from electronics to tool-and-die making to medicine to airplane flying or to engineering itself. These people may add up to a tenth of the male population, and a far smaller fraction of the women. To accept these fractions as inevitable is to accept an intolerable division of thought, to admit complacently that most Americans are living among shadows, albeit shadows cast by the latest luminous semiconductors. The world will not bear such a division, but the shadows seem to be growing every day, on the campus or in the shops outside the university filled with Zodiac and Tantric symbols. This is not to say that the mystical and the occult should or could be exorcised by rational argument alone; such a view is naïve, even destructive. But it is to say forcefully that the climate of interest in and the feeling of oneness with the scientific picture which underlies our world today is rapidly cooling.

Yet such a picture is indispensable for moderns. The world we inhabit, increasingly man-made, remains embedded in space and time, must be fashioned of matter and energy, can undergo change only by the interplay of energy and entropy. In many ways it is the physicists' world that people inhabit. And unless they understand it, they will dwell there with power they do not master and hopes that they are able only partly to fulfill.

Models of the world become fixed in the mind in the earliest years. Well-founded, based on tentative and testable schemes, they can persist and develop, with appropriate changes, for a lifetime. Insufficiently established, ignored, or actually made fearful, or derived from lore or from muddled perceptions, they can limit a person's entire lifetime and distort a whole society. This is the stake we have in teaching good physics in the schools.

Scientists sense out of their own deepest feelings how powerful is the fascination in some of our young with the world of phenomena. Recent studies have given us a measure of this power through its effect on schoolchildren.[4,5] Before the ninth grade, half of those children who as college students will concentrate in physics have opted for science; the figure reaches 95 percent prior to graduation from high school. If only to help them build their expectations soundly, we must promote the cause of good physics.

But good physics is not mere verbal formulas. It is not test answers or theorem results or knowledge of concepts as measured by verbal behavior. Not the word but the deed must be the watchword; it is not the names of the animals that matter—Adam knew the animals before he named them. Nor is the deed to be prescribed: we are not after dexterity with this or that instrument or the solution of this or that equation. All of the foregoing are surface phenomena; we must act *in depth.* This is the cutting edge. Depth is not mere cognition, though it must include cognition. It is also affect, manifest in attitude, interest, even delight. When fear and distaste accompany however letter-perfect a recitation of tables, units, terms, or laws, cognitive learning cannot take place.

3.1 THE IDEA OF PRODUCTIVITY

Linguists use the term productivity in a special way. They do not mean a ratio—output per man-hour—for theirs is a single measure, the extraordinary output of new sentences by men out of their stock of words and syntax. It is this that makes human language such a remarkably revealing mirror and so potent a tool of the mind and spirit. The same test is most congenial for any real understanding of mathematics and science at every level. What can the student do with what he knows to make a "new sentence"? Traditionally, what is tested is not the ability to make a new sentence but the ability to repeat a learned one, at best somewhat rephrased and at worst merely chosen from a set of alternatives, like the door-pecking of a pigeon.

Recently, the notion that education implies a potential for innovative ideas, even modest ones, has been given an anatomical base. An accident victim who had suffered severe brain damage was not once able, during nine years after the accident, to utter a new sentence. But she could repeat sentences told her, she could even cap familiar verses with their last lines; and she could learn the words and tunes of new songs—written after her injury—by listening to them played on a tape. At autopsy it was revealed that the brain areas for speech, both those that may be called executive (grammar, sounds) and those storing words, were intact, but their only connection to the rest of the brain was to their own auditory input and speech output channels. The experiences and recombinations made all over the brain were isolated from speech. This grim case is a kind of caricature of the speech and word sentence style of rote learning. We have spent much effort on training parrots. Humans so trained may often come to dislike and to flee from their tutors and their lessons. That is the price of such schooling.

Indeed, while symbols and the concepts behind them are the very fiber of physics, the nation today is inundated by symbols. TV and printing are ubiquitous; words and images flow in torrents. Once it was otherwise. Whatever their cultural and material deprivations, the Yankee farmer and his family had a richness of experience, from cow or churn or firewood or saw or moon or hepatica or gingerbread, from birth and dying. Symbols were in short supply; abstract knowledge was rare. The Bible was the TV set. Even the millhand in the industrial mill lived a life more of phenomena than of symbols. The *schools* taught the symbols, the three R's. Now the school has a new task. Phenomena are rare at a simple level. Animals and plants and sky and simple machines and home crafts have been moved out of everyday experience. Experience has become social dealing with others, or else it is symbolic, in print or in glowing electron target. What must the schools do? The answer seems clear. Provide guided productive experience with phenomena, as richly, widely, cheerfully as possible. The abstract world will enter, never mind. The schools need to cultivate affect and experience as much as theory, and even more!

3.2 GRADE-SCHOOL PHYSICS

It is certainly implicit in a curriculum whose outlook is on the world's phenomena that not very often should physics be singled out, away from astronomy, geology, biology, psychology. The tendency will be to merge, not to split. On a finer scale, of course, splitting will occur: The study of paper airplanes and bird flight is more physics than chemistry, or even biology.

Three such curricula for elementary school science have been implemented nationally in sizable efforts, mostly federally financed. They all direct themselves to both cognitive and affective learning; they all make use of material, not mere textbooks and pictures; they all merge the sciences rather than split them. But in degree and in detail they differ widely. The most nearly conventional curriculum* developed by a team of scientists and teachers under the leadership of a psychologist, emphasizes the conceptual growth of the student and directs each experiment rather tightly at some specific objective, be it an understanding of measurement or of animal feeding. Another,† led by a theoretical physicist, aims more, though by no means all, of its effort at an explicit introduction to some very general concepts, such as systems and interaction, or observers' viewpoints, and often finds its materials themselves becoming symbolic—arrows, or little men for observers, for example, or diagrammatic planes of one sort or another. The third curriculum,‡ unabashedly opportunistic, presents real materials, almost never symbolic, and demands much more of a research attitude in student, teacher, and classroom. The answers do not flow easily from real apparatus, as every designer knows. The paper written by a leading member of the originating team speaks of "messing about in science" (quoting Rat in *Wind in the Willows*) as an indispensable precondition to more formal and specific learning.

The context of the curriculum is of course the classroom. These three approaches (it is not implied that no other important ones exist) clearly suggest different classrooms. The first can make do with the present classroom, given equipment and a teacher willing to let the class do some work of its own. The second, more flexible, still places much responsibility on the teacher for guiding the students toward the right formulation. The third cannot really enter a scheduled, neat, and synchronized classroom at

* "Science, A Process Approach," American Association for the Advancement of Science; Xerox Corporation.

† Science Curriculum Improvement Study, University of California, Berkeley; Rand McNally.

‡ "Elementary Science Study," Educational Development Corporation, Newton, Mass.

all. Instead of depending on lesson plans, its mode is more that of research, where results are either delayed or come out unexpectedly early, where water spills and bulbs burn out, and the neat plan suggests two others before it is done.

That is the rub. Can classrooms approximate the spirit of research? In almost no way, certainly not in a school given over to discipline, to the leadership of the teacher, to the authority of the text. This conclusion is becoming clearer each day. But in a world where the complex and the symbolic are no longer the property of the schools, the schools cannot survive if they present one pattern alone for learning, the pattern of the lecture, the drill, the practice exercise. No single pattern will do; indeed, the drill and the lecture definitely have their place, and it is an important one, just as a measure of discipline must accompany most learning. What we need is a mixed strategy, and what school physics needs is a classroom—or perhaps better yet, a shop or lab out of the school—in which a mixed strategy based on phenomena can be employed. Some quiet and some noisy time; some book work and some work with water, pump, and weight; some niceness of attention and some sheer excitement of first discovery.

But what in fact are the new curricula able to accomplish? Are they on the right track? In answering these questions, we must bear in mind the old saying, "Experiments in education never fail," and seek standards that are objective and that probe deeply for evidence of the kind of understanding and performance we are advocating.

Our evaluation begins with a recent paper by McKinnon and Renner, "Are Colleges Concerned with Intellectual Development?"[6] The authors showed that three quarters of a group of incoming college students did not have reasonable command of primitive logical thought processes, a finding that instructors of introductory physics classes would generally agree with. This finding was based on the scheme for testing logical thought developed by the Swiss psychologist Jean Piaget.[7] McKinnon and Renner then ask: "What evidence exists to demonstrate that logical thought can be promoted among all levels of students?"

Renner and his associates are among those seeking such evidence. In one study,[8] they selected 60 first-grade children who had been exposed to inquiry-oriented science instruction and, as a control group, an equal number from a conventional program matched in response to readiness for first-grade work. Tests based on Piaget's analysis of the development of the notion of conservation were given at the beginning of the term and again after four months of school. The main feature of the test results was that the test group not only outscored the control group, but its gain was larger by 50 percent. In another study* they examined the performance of elementary pupils who had experienced science only through inquiry-oriented study for nearly five years of school in comparison with a matched control group. The two groups were tested in six areas of performance: observing, measuring, classifying, experimenting, interpreting, and predicting. In each of these categories, the first group was superior. The first group seemed, in addition, to be more creative in designing possible solutions to the tasks set and was less willing to give up.

If evidence of this kind accumulates, and so far there is little or none to contradict it, the nature of educational practice will be affected profoundly at all levels. We will be faced with the following hypotheses, adapted from McKinnon and Renner: (1) The educational experience does not now promote logical thinking in most students; (2) an abundance of inquiry-oriented courses, taught by teachers who themselves are products of such courses at the college and university level, must come into being in the schools before an alternative to the first hypothesis can be accepted.

3.3 THE SCHOOL TRIANGLE

The school is a complex structure—not merely a building, but a network of rules, usages, and plans which brings so many people together for so many hours, grouped roughly by age, and provides them with a desk and chalk, in poverty or in wealth. Whatever the structure, the school requires a special set of human relationships between teacher and students and among the students. Given the structural rules, the human circumstances, there remains the material content and nature of the school: its building and, above all, the equipment of the schoolroom. It is not easy to study anything well (and above all, no science) in a room whose desks are slant-topped and screwed to the floor and where nothing is at hand but a grudging supply of aging books and a little paper. It is the poverty of the school room, often evident even when the building cost is high, that stands out for many visitors who see elementary schools that work badly. A good schoolroom is rich in things, not necessarily because of the expenditure of budget funds, but because it abounds in invented, scrounged, lovingly sought out and made possessions—the paintings, maps, collections, pets, graphs, plants, tools, and trophies of the children who year by year form

* M. C. Weber and J. Renner, "How Effective is the SCIS Science Program?" (unpublished manuscript).

its shifting culture. A room given over to science cannot help but have such a look.

Structure, human relationships, material content—that is the school's triangle. And it is not alone the school's; no university physics department can manage without a good mix of these three categories. It is sad when young children do not enjoy the flexibility of structure, the mutual respect and willingness to learn, and the inexpensive tools for learning and experiment which even a modestly successful college physics department nearly always has. The school, of course, needs far simpler materials than do research workers at other levels; but it cannot survive extreme poverty. Books, paper, and pencils are not enough for the three R's; they are surely inadequate for science.

This is not novel, of course; many programs recognize the need. But the test is, again, not dollars spent but productivity. Hand tools, raw materials, basic facilities like outlets and sinks, some access to really grand possessions which may be shared, e.g., microscopes, are the test of such a provision. The prepackaged approach of the curriculum projects is an attractive start, but it is only a start. Neither research nor learning does well if all the equipment is provided ready to use.

Having some idea of the conditions favorable to the several new science curricula, we now ask, what is their likely fate? Do the schools have a place for them? Is it realistic to imagine a component of special interest to physicists? Who will teach the new curricula? These are difficult questions to answer, despite the existence of monumental compilations of statistical information.

Science Teaching in the Elementary Schools,[9] based on a survey made a decade ago, remains the most comprehensive source of information. The report maintains that the schools regard science teaching as an important responsibility in all grades. Most frequently, it is taught as a separate subject. The kindergartens give about an hour a week to science, the seventh and eighth grades in large schools three to four hours. The spread from school to school, depending mainly on size, is considerable, but a national average for all grades, weighted according to student numbers, is about 100 minutes per week. In the years from kindergarten through the eighth grade, this amounts to 450 hours. Our interest is primarily in the fraction of this time that can be allocated to physical science.

But the report ignores content, and little information is to be found anywhere on this subject, at least insofar as average practice is concerned. We must fall back on whatever indirect clues we can discover. One of these is the general conception of science in terms of three broad categories: physical, biological, and earth. A faint clue can be found in a list of equipment found in the average elementary classroom. About one third of the equipment is intended primarily for physical measurement (although no information is given about the frequency of its use). It may not be amiss, then, to claim about 150 hours for physical science in all of the elementary grades. Whether so much time is in fact spent for this pupose will be questioned by many, but our point is its availability.

In more than 80 percent of the schools, the classroom teacher alone handles science in grades up to and including the fifth; for higher grades this is true in 70 percent of the schools. Consultant help is available to more than 40 percent of the schools, but, of these, only 15 percent have an elementary science consultant and 29 percent, a classroom teacher with special competence in science. About half of the schools for which such help is available report using it rarely or never; when such consultants are used, it is most often to provide materials. Only 15 percent of all schools spend as much as $1.50 per pupil per year on science supplies and equipment; half spend less than 40 cents.

It is encouraging to learn that over three million grade school pupils, about 10 percent of the total, are engaged at present in the three largest new curricula.* But we can see the obstacles in the way of widespread adoption. The new curricula require somewhat more time, two to three hours per week over the first six grades. They demand about $10 per pupil per year, large compared with present expenditures for science but small compared with the national average total expenditure, which is about $800 per pupil per year. They demand teachers who have themselves learned in a mode that encourages inquiry.

3.4 THE TEACHER

The classroom teacher is in the best case a guide to the use of tools of mind and hand; a careful student of the children, sensing their needs as individuals; a source of some knowledge, not all; and a model of attitudes the children can use for growth. What the teacher must impart above all is a desire to learn and a means of learning. But the desire and the means need a subject matter, and that is the role of science, important in itself, and accessible by nature and by tradition, free of the trappings of family, culture, and social status. The experiment can work even if your accent is not what the teacher wants, and it can fail even though you are the favorite of the school and your whole family has done well. This honesty of the

* R. F. Tinker, Commission on College Physics, University of Maryland, private communication.

objective world is one strong point for science in general and physics in particular.

How can physics be brought to elementary school teachers? The needs of mankind and the opportunity of molding those who mold our young cry out that it is necessary to teach them in the way we wish them to teach, not by fiat and by rote, not by authority and by verbal formula, but with a sense of the phenomena and with some chance to make productive use of the concepts. These criteria seem more important than any dictum about the content itself. One could sacrifice the idea of energy or of atom quite cheerfully to see those qualities in their proper place. Few are the college courses that take such a position. This is perhaps the greatest single deficiency of physics in education. We do not honestly try to have the "terminal" physics student understand or like anything. We have preferred that he cover a lot. Never mind that he promptly forgets the coverage and takes away mainly a sense of inadequacy and confusion. What one gets is either a rigmarole of names and tricks or a large body of material that is like nothing so much as a new and less humane mythology—a myth of carbon cycle and nucleon number to replace the myths of Greece or Norway. If nothing else changes, this ought to change. One might or might not teach the young teachers literally with the materials intended for the children, but their exposure to science ought to be conceived in this spirit, and it would be far better to use the same materials than to follow the traditional route.

3.5 THREE ORDERS OF MAGNITUDE

One can see three time scales in learning within the realm of the school. A conventional course takes, say, 10^2 hours of attention. (We suppress all factors less than 10.) A lecture, or visiting a museum, or a trip to a factory, takes 10^0. There are not many models for the missing power of 10. It seems plausible that, in addition to the provision of many new forms of 1-hour study of parts of physics (or of science generally), adult students (high school seniors and up) should find many ways of expending 10 hours on a topic. Such a span is the sort of time it takes to acquire some small craft skill—beginning jewelry making or the like. The mode that begins to appeal, especially for eventual teachers, and in fact for all who encounter physics save those who will make a career of some engineering or scientific specialty, is one of making up courses by the union of many choices of 1-hour exposures with a small set of 10-hour concentrations. The 1-hour efforts—lecture, lab, demonstration, tryout, museum visit, theater, whatever they may be—are meant to establish familiarity and interest in some topic. Since there must be a wide choice, most of them will serve in fact only as part of an initial acquaintance with ideas or phenomena. Such is a visit to a good museum of science or art. But after many such introductions, one can expect a few deeper involvements: making something really work, following some real argument to its conclusion, searching out data from instrument or book, or the like. Several connected visits to a single complex place, with connective preparation, would sum to 10 hours. A number of 10-hour experiences (they might be 6 or 28 hours, of course) would make up what passes for a rather light term of college work. That hierarchical organization seems closer to the world of accomplishments than the highly uniform structure in which no gradation of importance serves to distinguish lectures on Newton's laws from those on calorimetry. One knows that not every topic, not every page of the text, is of equal importance. Indeed, importance varies from person to person, and from time to time for each person. So far, we recognize this very little, even in the typical college course.

3.6 INTERMEDIATE SCIENCE

Usually given in the ninth grade, a course in general science or in physical science is the last opportunity for reaching a very large proportion of our population. Seventy percent of all pupils passing through the system take this course, five times the number in high school physics. Girls populate it in equal proportion to boys, whereas in high school physics the proportion is only 1 to 3.[10] Perhaps 10 percent of the *teachers* of the course last studied physics when they were ninth grade pupils.

The strong strategic position of the intermediate course is often overlooked by casual critics of the educational scene. Pupils come to it at an age when enthusiasm is keen and mastery of performance and reasoning is developing rapidly. The external world is beginning to take shape. A science course at this level is a terminal experience for some; for others, it is a basis for the discipline-oriented studies of high school.

The strategic value of the ninth grade course has not been lost on curriculum designers and teachers. About twenty-five general science and physical science courses have been counted recently. Many are smaller, individual efforts. "Introductory Physical Science" was developed by a group* carrying on the work of the Physical Science Study Committee (PSSC), which had previously introduced a new physics curriculum for high schools.

* Introductory Physical Sciences (IPS) Group, Educational Services, Inc., Prentice-Hall, Inc., Englewood Cliffs, N.J.

"Intermediate Science Curriculum Study" comes from the Center for Research in Curricula for Science and Mathematics,* a resource development group in the Southeast. A number of courses have been supported in part by federal grants.

These courses share the intent of leading students through a systematic approach to a definite goal such as understanding matter and energy or the basis for an atomic model of matter. They aim to build the pupil's confidence in his own ability to answer questions by experimentation and observation. They mix activity and reading intimately. They face honestly the practical difficulties of even the simplest experimentation, which produces ambiguous results at times. One of the new courses introduces the caloric theory of heat in a serious way in order to enable pupils to discover that a reasonably satisfactory theory can be discarded when a better explanation for observation is found.† A common theme, stated in "Introductory Physical Science," is to make evident that "science does not deal with absolute truths." Another is stated: "When we 'explain' something, we build a bridge between the familiar and the unfamiliar by using mental pictures and theories and by making comparisons."† The mode is one of combining student initiative and discovery with guided progress, the teacher being urged to remain in the background.

The point has been made often and demonstrated innumerable times that science study has features that make it accessible to pupils with low reading ability and constricted cultural background. While the new curricula do emphasize nonreading activity, their textual material makes no apparent concessions, and from time to time logical gaps and verbalisms appear, requiring prior reading or other experience or, equally inconsistent with the new mode, assertion of authority. Special effort is required in order to match the curricula to the backgrounds of underprivileged students.

3.7 THE HIGH SCHOOL COURSES

Two well-known courses[11,12] dominate high school physics today. One of them, that of the PSSC, has been in use for a decade or more and has gained wide national, and even international, acceptance. "The Project Physics Course," having completed its incubation period, is just now beginning to

* Intermediate Science Curriculum Study, Florida State University; Silver Burdett Co., Morristown, N.J.

† *Physical Science, Laboratory Approach,* J. H. Mareau and E. W. Ledbetter, Addison-Wesley Publishing Company, Reading, Mass.

reach a substantial number of students. Each of these courses has been shaped by a main idea; PSSC strives to present a rich, tight line of inductive argument, whereas Project Physics unfolds its subject in the context of the intellectual history of man. Both were produced through the collaboration of college-level and high school instructors. Both accept the structural assumptions of the conventional high school—classes, books, tests, hours—although the newer course provides a set of options, as a step toward a more relaxed format. The target of both is the student who has academic life well in hand, although portions of Project Physics, which makes a genuine effort to be attractive to students who are not so strongly science-oriented, have been used successfully with disadvantaged students. For those more inclined toward practical concerns and applications, there is a third course, "The Man-Made World."[13] This course is based on engineering concepts such as feedback, stability, and optimization; it introduces students to mathematical modeling and to a simplified computer. Yet it has not met with a popular response.

In the light of the effort expended on these projects, the statistics on high school physics enrollment are distressing. The NSF reports[14] that "in every (science) course except physics a greater proportion of the students in the appropriate grade enrolled in the course in 1969–70 than in 1948–49. . . . [T]he proportion in physics declined steadily." The figures for physics in U.S. public high schools are: 1948–1949, 28 percent; 1960–1961, 23 percent; 1969–1970, 18 percent. These data are compared with those for other subjects and with total high school enrollment in Figure XIII.1.

The impact of Project Physics has yet to be registered in the above statistics. Its proponents believe that, through appeal to a distinctly broader segment of the student body, including girls, it will alter the trend in physics enrollments. However, if we have learned anything from the statistics, optimism in regard to the impact of any single new curriculum should be restrained.

A more striking example of the subtleties involved in producing educational change by design is difficult to imagine. The care that was exercised in the development of these new courses was immense. In them, concepts are developed on a sound basis of experiment and analysis. The unfolding of the material follows a clear line. The physics is as nearly impeccable as one could hope. Several versions were tried before final publication. All have been received enthusiastically. What is the problem?

Although no dogmatic answer will suffice, it is difficult to escape the conclusion that the desire for a tightly organized, logically constructed presentation tends to assume precedence over the realities of the classroom, that is to say, the nature of the learning process, the student's need for

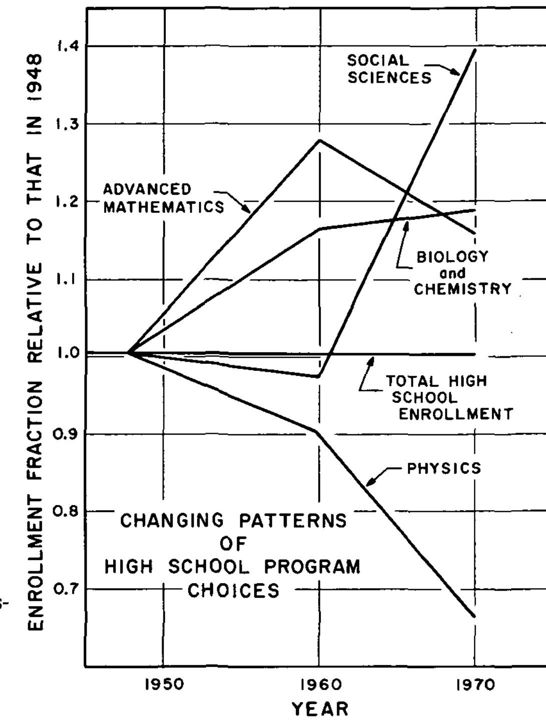


FIGURE XIII.1 Fractional enrollment in certain high school science courses relative to the fractional enrollment in those courses in 1949. Note that the drop in physics and the lower slope in biology and chemistry took place during the decade following the introduction of the new high-powered curricula. [Source: The National Science Foundation.]

repetition, and the teacher's need for breadth to enable him to handle the unexpected (for not even a committee can anticipate all of the confusions and misinterpretations of a student).

The material of PSSC, Project Physics, and Man-Made World will retain its validity and will be used for a long time. The issue is the manner in which it is used. Goethe has written: "Most thoughts have been thought already thousands of times; but to make them truly ours, we must think them over again honestly, till they take firm root in our personal experience." A model that is well adapted to learning in this spirit is that of the new elementary curricula.

In fact, in only a few years, high schools will be populated with students accustomed to the "inquiry mode" of learning. They will need a less linear, verbal, and formal course structure than is currently incorporated in any of the major high school curricula. The task for curriculum planners or teachers of high school physics is to adapt the excellent material now available to a freer, more flexible style of presentation and use.

This conclusion, that a change in teaching procedures is needed, is found also in a recent report by Elizabeth A. Wood.[15] Her survey of the problems of science teaching in the schools was based on interviews with students, teachers, and administrators. Anecdotal rather than statistical, it is not readily summarized. Dr. Wood offers recommendations that are pertinent to the present report. Many are as valid for the lower grades as for high schools. We quote a few that reinforce the attitudes expressed here:

> We must begin to plan now for part-of-the-day teacherless teaching in the schools. This involves not only better teacherless materials, but also attitude reorientation on the part of school officials, parents, teachers and especially pupils.
>
> Since projects undertaken by bright ambitious students soon get beyond the stage at which the teacher is able to give needed advice, it would be desirable to have a panel of local scientists in colleges or industrial laboratories who were willing to act as advisors to high school students.
>
> Materials are needed for the students for whom reading and writing constitute a formidable barrier to the learning experience. In science we have an opportunity to "hook" them on the fun of learning through experimental work, with the exciting possibility of motivating them to read as a result.
>
> Information covering science requirements for college entrance for a representative sample of colleges should be printed in concise form and made available to the schools. Misconceptions about college requirements inhibit innovation.
>
> Opportunities should be sought for sympathetic cooperation between the Education Community and Science Community.

Finally, Dr. Wood sees "an increasing divergence between the level of science in the text books (in terms of degree of abstraction, expected background, vocabulary and general sophistication of approach) and the level of competence of the *average* student and the *average* teacher to cope with a science course."

If the intent of high school science is general education rather than the generation of specialists, many ask, why should not chemistry and physics be taught in combination? Is there not a core of subject matter common to both, and is not this core the principal object of attention in both subjects as presently taught? Is it not because of tradition and vested interest that a pattern a century or more old still continues? And if a suitable combined course were to be introduced, might it not arrest the disturbing trend in physics enrollment? These are cogent questions. Temperature, pressure, atoms, molecules, energy—the list of common topics is indeed lengthy. Both subjects are quantitative, both are laboratory sciences. Differences in research conducted today in physics or chemistry departments often depend more on local history than logical distinctions.

However, the two subjects are not identical. A report on education in chemistry[16] states that "a reasonable short definition of the scope of chemistry is: Chemistry is the integrated study of the preparation, properties, structure, and reactions of the chemical elements and their compounds, and of the systems which they form." A high school textbook[11] explains that "physics . . . deals with such features of the world as time, space, motion, matter, electricity, light and radiation; and some features of every event that occurs in the natural world can be seen in these terms." Incomplete as any such brief statements must be, they do convey the impression of distinctly different interests and approaches. The questions then reduce to whether there is pedagogic merit in bringing together two disciplines that in a certain basic sense are different, though they have much common subject matter. For a course designer, it becomes a matter of finding a natural progression or alternation between the basic material of physics, that is, space, time, motion, force, energy, light, electricity, and magnetism, and that of chemistry, which is elements, compounds, reactions, synthesis, and analysis. From the standpoint of conventional physics, these subjects of major concern to the chemist occur very late in the progression of manifestations of basic physical notions and laws. If they are not to be so treated, then a carefully designed sequence of topics is necessary so that the course will not consist of a collection of inhomogeneous lumps of subject matter. Until recently, the principal efforts at a combined approach have been directed toward the college level; at least one of these[17] shows promise of adaptability for use in secondary schools, if teachers are prepared to move back and forth easily between the two disciplines.

A plea for a more thoroughgoing rationalization of high school science has been offered recently by Swartz, who proposes the abolition of the standard sequence: earth science, biology, chemistry, and physics.[18] This sequence is precisely the inverse of the scientific relationships among them. Swartz points out that few students notice the irrationality because of the rapid decline in enrollment through the progression.

Swartz maintains that the way to break out of the conventional sequence is, in keeping with a theme that pervades this chapter and later ones, the introduction of a modular curriculum. He complains that there is, at present, no way to sample physics or to study one particular topic that might interest a student at a particular moment. "To find out about rainbows it is necessary to solve problems about boats heading east at five miles an hour crossing streams flowing south at two miles an hour."

An attempt to break out of the traditional pattern, at least in physics, has been described by Deall.[19] He has divided the standard year-long physics course into a large number of small steps and makes these available to students over a four-year period. Beginning with the ninth grade, a student comes to the physics room during free periods and works through the course at his own pace. He receives a quarter of a standard unit of credit for each year's work.

It is possible to envision a unified approach to science, carefully articulated so as to introduce basic topics and their ramifications throughout science, with emphasis gradually shifting from the simple systems of physics to the complex ones of ecology but always in contact with all of the fields. We are far from being able to realize this ideal, but we must encourage and assist those who are willing to take steps toward it.

3.8 GUIDANCE

Students need counseling and encouragement. Those in the schools who provide these services should understand the kind of appeal made by a subject like physics, the demands it makes, and the opportunities it offers, not only within its own disciplinary confines but also as a basis for other subjects. They must begin to understand the opportunities that exist for women and for members of minority groups. Earle Shaw* has given a graphic description of his difficulties as a black student, being given advice in high school and college that needlessly attempted to discourage him in his route to a PhD degree in physics. The physics profession, through the American Association of Physics Teachers or the American Institute of Physics, should arrange that suitable background material is made available to high school and college counselors, as well as to teachers.

Physics continues because talented young people find it challenging and rewarding to study. Some who begin the study of physics in college find it not to their taste or talent.[20] The evidence that decisions respecting scientific careers are made by significant numbers of children prior to high school[4] is strong reason for making an effort to identify talented youth in the lower grades and provide them with opportunity to explore science more adequately than is possible now. The object of such a program should be neither to force early decisions nor to seduce talent away from other fields. Its purpose should be to identify talented children and help them begin to sort out their attitudes toward science on the basis of better information than is now made available to them during their formative years, when their contacts with science are slight and with scientists, vicarious.

The National Science Foundation has been operating a program for talented students in high school and in college, and we note with approval that it is to be continued. At the same time, we suggest that consideration be given to a suitable extension into lower grades.

* "A Black Educational Experience," paper presented in the Summer 1969 meeting of the American Association of Physics Teachers.

4 Teaching the Teachers of Science

Teachers will teach in the way they have been taught.

OLD CLICHÉ

In the face of changing patterns of physics in the schools, there is no point in forcing more new teachers through a system so demonstrably mistuned to the new modes. To set it right will require some profound changes in the attitudes of a variety of sectors involved with education. Basically new concepts embracing re-evaluations of needs, opportunities, and potentialities are needed. Such changes will not take place in accord with a grand scheme, nor will they all occur simultaneously. Where initiative is possible, it must be exercised. There is a point of entry into the system for college and university physicists. They need only to comprehend the problem, to realize what is at stake, and to muster their determination and their patience. They must undertake to teach the teachers.

Before explaining what we mean by this declaration, we give a brief overview of the present corps of teachers, focusing on aspects of their backgrounds and activities that we regard as pertinent to our critique.

4.1 THOSE WHO TEACH PHYSICS IN THE SCHOOLS

In 1970, the elementary classrooms contained 1.2 million teachers, the secondary schools about 1 million.[10] Apart from growth and changes in the pupil–teacher ratio, the turnover rate of teachers has been about 8 percent per year. The U.S. Office of Education estimates[21] that between now and 1980 a small decrease in elementary enrollment and teaching staff will compensate for increases in the secondary grades, the consequence being that the total number of pupils and teachers will remain nearly constant, as will the turnover rate. The total expenditures by elementary and secondary schools in 1970 amounted to nearly $50 billion.[10] The salaries of classroom teachers constitute about $20 billion

Despite the existence of much detailed information about teaching practices in elementary schools, the subject matter preparation of the teachers does not appear to have been a prominent item for study. The most recent report from the U.S. Office of Education on science teaching in the elementary schools does not take up the matter of teacher preparation. In the absence of direct information, we turned to two indirect sources. The American Institute of Physics[22] estimated the average requirement in physical science imposed by fifty state universities on elementary education majors in 1961 as 2.9 semester hours. A second source was the records of 247 students in one university who completed requirements for an elementary provisional teaching certificate in June 1971. Seven percent listed a physics course and 8.5 percent, a chemistry course. The remainder satisfied the science requirement through descriptive courses in biology and the earth sciences.*

The subject matter preparation of secondary school teachers of general science is presented in Figure XIII.2, which summarizes data obtained in a recent survey by the NSF.[23]† Some uncertainty exists because the NSF did not collect data specifically on general science; instead, the diffuse category, "General Science and Other," was used. In addition, there was a large fraction of nonrespondents to the survey. The surmise that the data do not mask a substantial background in physical science is supported by the similarity between the distributions shown by the two extreme sets of bars in Figure XIII.2. In an earlier survey,[24] the NSF ignored the physical science preparation of general science teachers. The conclusion that general science teachers are typically deficient in physics, and to a lesser extent in chemistry, at least as indicated by their college training, is clearly borne out by these data. About two thirds of the U.S. public junior high schools offer in-service training of some type for their science teachers, and about 80 percent provide some kind of specialist assistance. However, the subject matter distribution of these modes of assistance was not examined in the survey.

The 1969 NSF survey shows that no more than 1 percent of those teachers whose major teaching assignment was general science reported a bachelor's degree in physics, and a like fraction reported a master's degree in this subject; significant fractions were found in biology and general science and small fractions in chemistry and mathematics. A somewhat paradoxical finding, albeit hardly a comforting one, was that 19 percent of these teachers reported graduate credits in physics (the same percentage as for mathematics; 21 percent had graduate credits in chemistry; and 33 percent, in biology). Finally, a closer inspection of the data shows that 8 percent of the teachers had *no* college physics; for these teachers, their own exposure as pupils to general science probably constituted their only preparation for teaching the physics component of this course.

In view of the strategic position occupied by teachers of general science and its subsidiaries, physical sciences, biological sciences, and earth sciences, we suggest that in future surveys the NSF report unambiguously on the preparation of those teachers and on the enrollments in these subjects.

* R. C. Salyer, College of Education, University of Washington, private communication.

† We are grateful to J. C. Lewis and J. H. Andrews for additional data not included in this report.

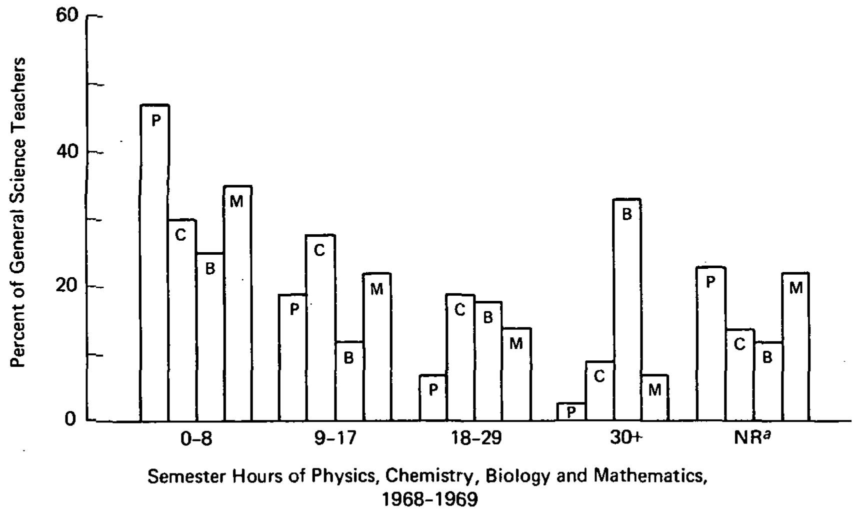


FIGURE XIII.2 Preparation of general science teachers in basic sciences and mathematics. [a]NR = no response.

The preparation of those who teach physics in secondary schools, measured by physics courses taken, apparently has improved, as can be seen in Figures XIII.3(A) and XIII.3(B). A detail omitted, because the comparison was not published, is that 2 percent of the 1968–1969 sample reported no credits in physics. Virtually all of the teachers hold a bachelor's degree (not necessarily as highest degree), of which about 20 percent are in mathematics, 17 percent in physics, 14 percent in chemistry, and the remainder largely in general science and the other natural sciences. Only 4 percent have a bachelor's degree in education. About 10 percent hold a master's degree in physics, and about 13 percent, a master's degree in other physical sciences or in general science.

Undoubtedly, because of the relatively small enrollment in physics, the assignment of teachers to this subject remains more haphazard than in the case of the other sciences; 59 percent of those teaching physics reported that their major teaching assignments were in other fields. This figure has not changed appreciably since the 1961 survey. The corresponding figures for other fields are: mathematics, 20 percent; biology, 25 percent; chemistry, 40 percent; earth sciences, 40 percent; general science, 40 percent. The 41 percent of those teaching physics who reported that physics was their main teaching assignment have a substantially stronger background than the average physics teacher. However, 9 percent of this group had no

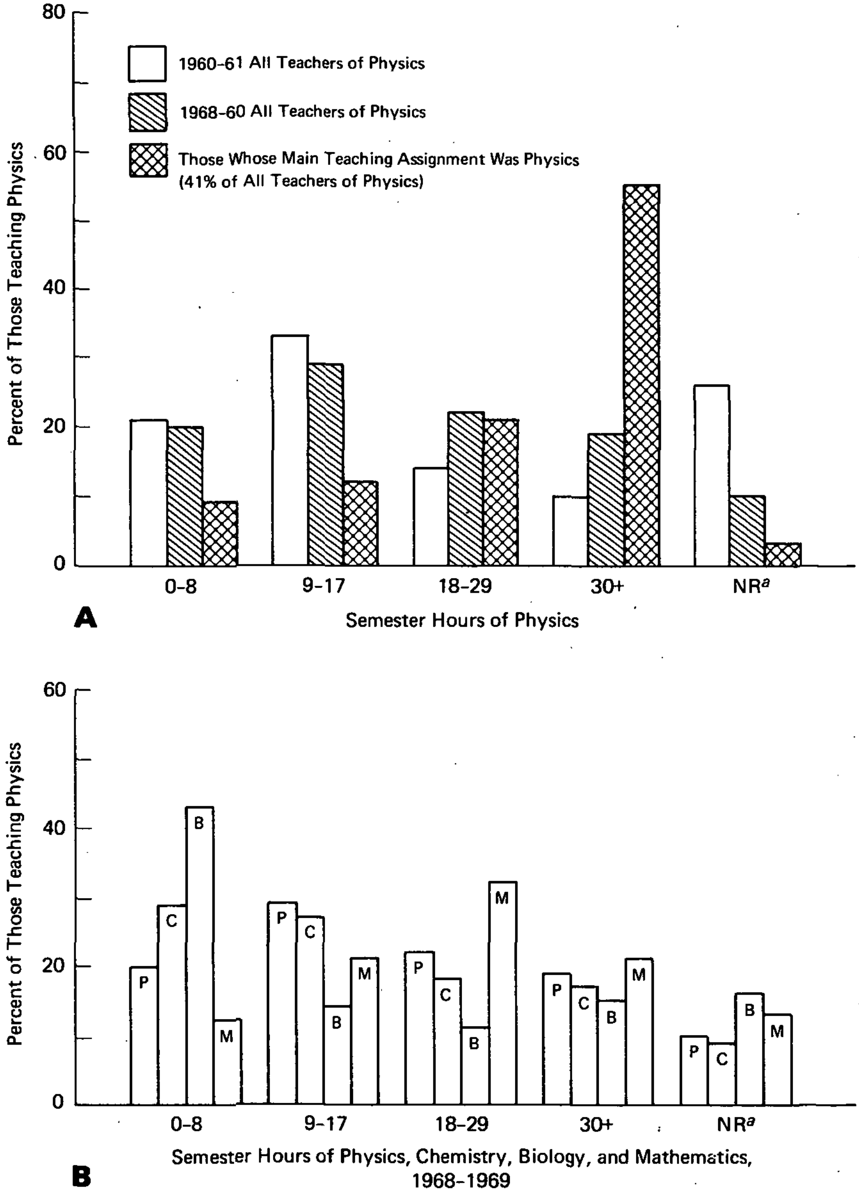


FIGURE XIII.3 (A) Physics preparation of all teachers of physics in 1960–1961, all teachers of physics in 1968–1969, and those, in 1968–1969, whose main teaching assignment was physics. (B) Preparation of high school teachers of physics in basic sciences and mathematics in 1968–1969. [a]NR = no response.

more than a year of college physics, and this subject, the report indicates, was their weakest science field.

We note that in conformity with earlier surveys and enrollments in the different science courses, the teaching of physics as a major assignment is confined to a small fraction, 3 percent, of all secondary school science teachers. This fraction should be contrasted with those for biology, 13 percent; chemistry, 6 percent; and mathematics, 38 percent. Like physics, earth sciences teaching is the major assignment of 3 percent. These figures have a special significance for those who plan and conduct the preparation of secondary school science teachers: No matter how well thought out, a teaching *major* in physics can be a productive enterprise in only a few special locations.

For teachers in service, there has long been the expectation that a variety of supplemental training programs would be effective measures for refreshing their acquaintance with the subject matter they teach. These programs take the form of summer institutes, in-service courses, academic year institutes, and the like. Many agencies, private and public, have provided support for all such activities. The NSF has been by far the most active (often under the prodding of the Congress).

Each year since the mid-1950's, about 15 percent of the nation's high school science teachers have participated in an institute; the NSF has supplied more than half a million stipends for this purpose.[25] The Foundation estimates that 40 percent of the present secondary school teachers have attended an NSF-sponsored formal program and that another 20 percent have participated in other programs.[23] Very little evaluation of the effectiveness of supplemental programs has been published, except for anecdotal accounts collected by questioning participants. One published investigation[26] concluded, on the basis of three standard tests given prior to and on the last day of four summer institutes for high school teachers, that participation had produced a gain in subject matter mastery and in understanding science as a process. Quantitatively, however, the gains in mean scores for the three tests were 13 percent, 6 percent, and 2 percent—so slight that it is doubtful that any long-term effects exist. There also is considerable anecdotal evidence to support the view that summer institutes are often presented at the same breakneck speed that contributes to the necessity for them in the first place. It is not surprising, therefore, that the executive branch of the federal government has sought for several years to curtail expenditure of this kind.

Judgments about the future of the summer institute programs must be made carefully, however. It will always be necessary for science teachers at whatever level to update their information and teaching procedures from time to time. Moreover, summer institutes have been the principal way to disseminate the new curricula and methods among in-service teachers; in fact, no other means exist currently. The interest of the federal government in science education, which led to the development of these new curricula, must extend to their dissemination. To be effective, institutes for this purpose must be monitored carefully to make certain that they follow and accurately transmit the new approach and procedures. The support program, therefore, must contain provision for some kind of control and evaluation.

We can point to nations in which regular, required, subsidized in-service programs exist. The attitude of the Soviet Union, for example, toward the need for continuing retraining was described recently in these terms[27]:

> It is well known that the human brain has a property to forget things, and so of course teachers working in schools will lose after a time the knowledge that they gained at the pedagogical institutes. So the problem is to recover this knowledge, and for this purpose there are about 180 retraining institutions. At these pedagogical institutions we free the teachers in June, give them their salaries, and require them to attend courses to raise the level of their specialties. This in-service training is compulsory, and every teacher has to attend a course every five years; otherwise they may not be allowed to teach.
>
> Recently we introduced a new program in which we laid down new ideas about how science has been developing recently. We teachers are, by our nature, slightly conservative. We are fond of those methods to which we are accustomed. Therefore the teachers did not welcome the introduction of these new programs. This somewhat handicapped the introduction of the program. To make the teachers able to transmit the new ideas to the pupils, we had to give them the new methods, and therefore during the change to the new method all of our teachers had to attend a one-month course to learn the new program.

Undoubtedly, supplemental programs have been of value; they could be more fruitful if conceived and carried out with greater comprehension of the problem they are supposed to alleviate. However, they do not constitute an easy solution to the problem.

4.2 BENEATH THE SURFACE

Relevant and intriguing though they may be, numerical statements about the courses teachers have taken, such as those quoted in the preceding section, do not assess the quality of either preparation or performance. Over a period of years, many surveys of teacher characteristics, in addition to asking about course credits, have included questions on attitude in an attempt to elicit factors important to successful teaching; unfortunately, these questions often suggest the desired responses. The staffs of summer institutes and others actively engaged in attempting to influence the attitudes and behavior of the teachers gain a different insight into the knowledge and attitudes the teachers actually carry with them in daily life. It is interesting to juxtapose two conclusions, one from each source:

> The science teachers have a strong desire for self-improvement; they have participated extensively in NSF and other educational training programs; they have a good subject-matter foundation in their backgrounds which they value; and generally they felt well-prepared to teach their subjects and were committed to continuing their careers in education.[23]

> We at Harvard Project Physics are already encountering great difficulty with teachers who have no background in the history of science, in the philosophy of science, and have a confused idea about the nature of the laws of technology and science. The teachers we now have in the high schools are the products of the educational system which we have been running.[28]

The second of these quotations is consistent with an observation made a decade ago by Gruber[29] who tested a group of 202 high school science teachers by asking each to make an outline, on some science subject he knew well, that would be suitable for a 20- to 30-minute talk by a high school senior. Gruber analyzed these outlines to determine the extent to which they showed perception of science as a way of studying, of investigating, and of viewing things; the degree of attention to given theory, uncertainty, and development leading to information and comparison of ideas or approaches; and the extent to which they incorporated informational aspects into the process-of-science framework. He found that more than 60 percent of the responses gave negligible emphasis, and another 10 percent moderate emphasis, to this attitude toward science. His finding is the more striking because each teacher had been selected as a Fellow to participate in one of a number of academic-year institutes intended to focus on the subject matter of science and mathematics. On the basis of this and other information, he concluded that: "High school teachers generally approach science teaching as a matter of conveying science as established facts and doctrines."

Such frankness is rare in the literature, but summer institute directors frequently are willing sources of unpublished opinion on the disparity between the transcripts of the institute participants and their attitudes and performance.

It comes as no surprise that elementary school teachers have even less understanding of science as an investigative and rational process than high school teachers. Torrance[30] found in a survey of a group of 1000 elementary school teachers that only 1.4 percent regarded independent and critical thinking as the most important educational objectives.

We conclude from these somewhat fragmentary pieces of evidence that teacher attitudes toward science and preparation for teaching science are seriously short of what is desirable. This statement is not a criticism of the teachers; it is the system that is at fault. Almost invariably, when asked what kind of additional study is most valuable, teachers ask for subject-matter courses. It might be said that they are seeking courses in methods, not the usual kind of course offered in departments of education, but courses that give them experience with the methods of doing science instead of aiming solely at efficient transfer of subject matter.

A fair question is whether, through courses of any kind, teachers can be induced to improve their understanding of science and alter their performance. Results of studies are beginning to appear, suggesting that significant changes in teaching performance occur after the teacher has been in an inquiry-centered course. Work at the Science Education Center of the University of Oklahoma exemplifies the evidence for this conclusion. In one study,[31] 30 classes of elementary school children were observed. Half of these classes were taught by teachers who had been trained for the use of inquiry-oriented science material, the remainder by teachers who had had no contact with any of the new science curricula. The kinds of science experiences given the students by the two sets of teachers, and the questions they asked, were compared statistically. In brief, the results showed that the inquiry-trained teachers gave their pupils more than twice as many science experiences as the others and asked 50 percent more questions. Moreover, the favorable ratios were much greater in the categories of experience involving measurement and prediction and the categories of questions requiring skill, analysis, and synthesis. The non-inquiry-trained teachers asked more questions in the categories of recognition and recall. Another study[8] investigated the extent to which elementary school teachers altered their patterns of reading instruction as a consequence of a summer workshop in "new science." This investigation also was based on an analysis of types of questions about reading material asked by a group of inquiry-trained teachers and a matched group of conventionally trained elementary teachers in the second and fourth grades. For each grade, the two groups used the same reading material. The investigation showed a significant shift to questioning aimed at levels of thought above recognition and recall, requiring translation into other situations, analysis, and synthesis. Questions eliciting opinion and attitude were significantly more frequent among the inquiry-trained teachers. A study of before–after observations of questioning in social studies classes gave similar results.

4.3 PUTTING A VALUE ON SCHOOL PHYSICS

Can we assign a value to the importance of school physics to society, or to physicists for that matter? "Importance" is a matter of individual qualitative judgment, but in attempting to assess it, we often seek quantitative comparisons for help. A suggested comparison sets the yearly cost of school physics against other costs that are more commonly associated with physics.

We begin by estimating the amount expended each year in the United States for teaching physics in the schools. We restrict our estimate to the salaries of the teachers, the major component of this expenditure. It is necessary to estimate the fraction of time a teacher devotes to topics that can be reasonably considered physics. Especially in the lower grades, this fraction is likely to depend more on the background of the teacher and on the facilities actually at hand than on any "adopted" curriculum or syllabus prepared by a central school administration. Accordingly, little information about the average fraction of time actually used for science, much less physics, is available. However, using published data and spot checking against situations in individual locales, one can produce a conservative estimate, which is the basis for Table XIII.1. The total given in this table may be inaccurate by as much as 25 percent, but only a rough approximation is needed.

The question before us is whether this total is commensurate with other sums for which the physicist customarily acknowledges responsibility. One such amount surely ought to be his own salary. Table XIII.2 gives some estimates for comparison.

Another kind of expenditure, and surely one for which the physicist eagerly exercises responsibility, is that for his research. Table XIII.3 gives current federal estimates of obligations for various components of research in 1971.

Comparing the various entries in Tables XIII.2 and XIII.3 with the total given in Table XIII.1 suggests that the nation is demonstrating as much willingness to support physics in the classroom as in the research laboratory. Let us suppose for a moment that U.S. physicists were offered an additional \$300 million per year. In a sense, they *have* been offered this sum, and it is now, and will continue to be, offered; the only "string" is that it must be used for educating the public, whose political actions will reinforce or hinder an enlightened policy for basic physics in the future. Thus far, physicists have shown little interest in the use of this large resource. A major challenge facing the physics community is to recognize this responsibility and develop means of exercising it meaningfully and wisely.

TABLE XIII.1 Annual Expenditures by Schools for Teaching of Physics[a]

School Level	No. of Teachers (thousands)	% of Time for Physics	Average Salary (1970) ($thousands)	Expenditure ($millions)
Elementary (Kindergarten–Grade 6)	1030	1.5	9.0	130
Intermediate (general science)	52	25.0	9.5	120
High school physics	21	32.0	9.5	64
Estimated total salaries of schoolteachers attributable to physics				310

[a] Sources: *Science Teaching in Public Junior High School; Science Teaching in Elementary Schools; Secondary School Science Teachers, 1969; Estimates of School Statistics, 1970-1971* (Research Division, National Education Association, Washington, D.C., 1970); and *U.S. Registry of Junior and Senior High School Teaching Personnel in Science, Mathematics, and the Social Sciences* (National Science Teachers Association, Washington, D.C., 1971).

4.4 IMPLEMENTING WHAT HAS BEEN DEFINED

Numerous observers report empirical evidence supporting the assertion that if an individual has not passed through a given learning experience or concept formation in science when he is a child and comes to the learning problem *de novo* as an adult, he goes through the experience at the same pace, encountering the same hurdles and obstacles, as does the child. The assertion has been documented in the case of future elementary school teachers, who usually come to their college science work completely innocent of any understanding of scientific concepts or models. Their progress is, if anything, slowed down by the confused memory of names and formulas they have been forced to learn in preceding years and that they proceed to utter in stochastic sequence, hoping against hope that a sufficiently meaningful juxtaposition will occur frequently enough to give them some "partial credit."

TABLE XIII.2 Salaries of Physicists, 1970[a]

Physicists	Number (thousands)	Median salary ($thousands)	Expenditure ($millions)
College and university faculty physicists	8.7	14.0	123
PhD physicists	16.0	17.4	278
All physicists	36.3	16.0	581

[a] Sources: Statistical Data Panel of the Physics Survey Committee; *Physics Manpower and Educational Statistics* (AIP-1969).

TABLE XIII.3 Federal Obligations for Research in Physics, 1971[a]

Source of support	Estimate ($millions)
NSF total support for basic research	38
AEC total support for basic research	190
Federal obligations for basic research	420

[a] Source: *Federal Funds for Research, Development, and Other Scientific Activities* (NSF 70-38).

The learning models we advocate are hardly new. Implicitly and explicitly, they have been defined by every philosopher of education from Socrates, through Montaigne, Rousseau, and Whitehead, to contemporary thinkers and researchers. We must implement far more widely what has been so long defined. We plead for approaches that invite future teachers to examine the questions of "How do we know . . . ?," "On what evidence do we accept . . . ?"–approaches such as those operationally exemplified by the elementary science materials described in Chapter 3.

An educationally viable way of fostering this attitude is to lead future teachers into individualized, self-paced, experiment-and-observation-oriented study of a few of the major concepts and general themes that cut across all of the new elementary science curricula (see, for example, Arons[32]). To this end one might draw from elements such as the following:

Formation of the concept of "property" of a material object.

The way in which a unique constellation of properties ultimately defines what we mean by a "material substance."

The quantification of some properties (such as density, solubility, thermal expansion, melting point) and the practice in arithmetical reasoning that this involves.

Elementary understanding of the description of motion and the law of inertia. (The vast majority of our college-educated population, not to speak of others, is completely Aristotelian as far as any comprehension of the concept of force and the law of inertia is concerned.)

Direct observation and correlation of the most elementary and apparent astronomical phenomena. (Too many college students are completely unaware that the stars have a diurnal motion.)

The capacity to discuss evidence for geocentric versus heliocentric models of the solar system, to leave this discussion *open* with children instead of closing it by assertion, and eventually to develop an understanding of the geocentric model in terms of the law of inertia.

Evidence for discreteness in the structure of matter, why we accept and how we elaborate the atomic–molecular model.

Comprehending "energy," "electrical charge," and "magnetism" as constructs with explanatory and predictive power in common, everyday phenomena and not as esoteric, material substances.

Having personal awareness of at least several bits of evidence that indicate a deep connection between electricity and the structure of matter (without involvement in the end results of modern quantum chemistry).

Understanding the distinction between heat and temperature and visualizing thermal phenomena in terms of the microscopic model.

Recognizing the meaning of the term "vacuum" and comprehending some of the familiar properties of air and of liquids.

Having some direct experience with the contrast between particle motion and wave motion, comprehending sound as wave motion that transcends our visual experience, and performing some of the experiments that lead to a wave model of the nature of light.

Taught in the manner we advocate, the above examples may easily occupy an entire year. If no time is left for strange particles, relativity, and the uncertainty principle, we should not mourn. More important are roots in genuine understanding.

The general principles enunciated above, intended especially for elementary school teachers, apply equally well to the preparation of high school teachers. We also must teach them in the manner we hope they will subsequently use in their own classrooms. The question of preparation of secondary school physics teachers has been studied by the Commission on College Physics (Panel on the Preparation of Physics Teachers). The reader is referred to a report of this panel[33] for relevant statistical information and for a detailed discussion of possible curricular structures and options. An interesting description of a working program has recently been given by Little,[34] who pays special attention to the ninth grade science teacher.

These analyses emphasize the importance to the prospective teacher of spreading his college program into areas such as mathematics, chemistry, astronomy, geophysics, and biophysics, rather than concentrating in a conventional physics major. A carefully planned physics minor, cognate to a major in another science or in mathematics, might be a realistic source of secondary school teachers. A perfectly reasonable level of subject matter coverage would be that which develops thorough understanding of the actual content of secondary school courses such as PSSC Physics, Project Physics, and Chem Study. Achieving just this objective would put us very far ahead of where we are right now.

4.5 FOR WANT OF A NAIL

The old allegory, it has been said recently,[35] describes the dilemma in regard to education. The new curricula offer promise of an eventual improvement in public understanding. Some believe they can help to motivate the culturally deprived. They are modest in terms of equipment and demands on physical facilities. Perhaps not ideal as they stand, they are loosening the tight conception of school science held in an earlier day and are stimulating further modification, adaptation, and evolution. Yet they are unlikely to have the effects wanted by their originators, their financial backers, and those who sat in the stands and cheered. This bleak opinion results from evidence that the *way* in which the teachers are learning science, more than the *amount* of subject matter to which they are exposed, determines the manner in which they use the new materials. McKinnon and Renner,[6] in their study and analysis of intellectual development, conclude that "secondary and elementary teachers do not take advantage of inquiry-oriented techniques so necessary to the development of logical thought, because college professors do not provide examples of inquiry-oriented courses."

We must acknowledge that this conclusion is essentially correct. The example set for most students in physics courses (those students with strong *a priori* orientation toward physics being perhaps less susceptible) suggests strongly that this subject can be conveyed by lecturing, with occasional demonstrations (prearranged), and, when a laboratory accompanies the introductory course, by predigested "experiments." The students are only learning to repeat "old sentences" that they soon forget.

There are traditional ways for science faculties to rationalize their historic indifference: Future teachers are in the hands of educators who load their programs with trivia; they cannot think quantitatively; they are not up to "honest" physics courses; the problems of elementary school science are so large they can be managed only through specialist teachers; the faculties of institutions without teacher education programs have no responsibility. All these are stated as reasons for perpetuating the existing situation.

Let us face the matter more squarely. Nearly all future teachers are required to study science in a subject-oriented department. An "honest" course is one that induces students to learn for themselves. Even granted a large increase in the availability of science consultants, the classroom teacher will continue to be the dominant influence in the classroom; no one is making a strenuous effort to produce science consultants, or even to use those who could become available. Physics courses desirable for the purpose of teacher education can serve admirably as liberal arts courses; thus virtually all physics faculties could usefully engage in their development. Given an 8 percent turnover rate of teachers,[21] a substantial impact can be made in a decade.

It is apparent to those few physicists who have tried teaching in the inquiry mode that college students intending to become teachers can be led to explore, measure, compute, err and recover, and draw unanticipated conclusions. In going through this process, they recapitulate the steps and require the same time characteristic of schoolchildren, clearly not because of any fault of their own. In essence, here is the problem, and also a promising step toward solution. There are precedents for any physicist who chooses to look and wishes to act.*

Ultimately, the responsibility for the quality of physics teaching in the schools returns to the college and university faculties. They are the custodians of physics, and they teach—or do not teach—those who emerge from their halls to teach the young.

5 Clients by the Quarter Million

About 250,000 college students enroll each year in introductory physics courses. It is impossible to know how many of these students plan careers in physics, because few are required to declare their intent in the early years of college. We do know that by their junior year, about 7000 have identified themselves as physics majors. It is a safe assumption, then, that 95 percent of the individuals in introductory courses hope to gain something from physics other than a profession. We nurture the smaller group with special formulas to develop the insights and abilities we think important for physicists; Chapter 7 is devoted to our efforts with them. In this chapter and in Chapter 6, we examine our last formal opportunity to make contact with more diverse components of society.

The students we meet in introductory college courses are so diverse that long ago we found it unworkable to treat them all in the same way. Great differences are apparent in their motivations, aims, and needs, not to mention the facility with which they communicate in mathematical language. Whenever possible, a college will offer differentiated approaches to physics that cater to different categories of students. In the conventional language of academic life, these categoreis are the general liberal arts students, stu-

* For a cautiously optimistic view of the trend in secondary school physics teacher preparation, see Strassenberg.[36]

dents majoring in sciences other than physics, premedical students, and engineering students.

Although the differences in the career aims, needs, and motivations of these students are apparent at once, further consideration suggests that there are important *common* features that should determine the kind of exposure to physics they receive. These features are sufficiently important to warrant general discussion before the distinctions among the groups are considered.

5.1 VALUES OF PHYSICS IN EDUCATION

What motives have we for improving physics instruction for college students who do not intend to become professionals in the field, beyond our natural enthusiasm to talk about ourselves and our work?

While it is as easy to defend the place of physics in a curriculum as it is to defend that of any other "liberal art," it might be well to see if there are any unique values to be derived from the study of physics, in terms of intellectual effort and discipline, that would confer special benefits on all the various categories of students.

First, since physics is an "exact" science in the sense that most questions, if properly put, have a verifiable answer, a student necessarily develops an appreciation of how problems can be solved in a systematic way. The difference between physics, on the one hand, and mathematics and logic, on the other, is that physics proceeds with less formality and rigor and is more readily related to everyday experience and the environment in which we live.

The style of physics, which is to break down complex situations into smaller, more manageable pieces, illustrates a technique of considerable applicability that can promote rational qualitative and quantitative analysis and thereby expose humbug.

The opportunity that physics provides for the understanding of natural phenomena adds to the enjoyment of life, not only for those who are curious, but also in the sense that this understanding reduces man's empiricism, superstition, and insecurity in his interaction with the rest of the universe.

Recognition of the role that human intelligence and judgment have played in formulating the principles of physics, the role that taste has played in their statement in elegant form, and the applicability of these principles to vastly different situations is surely worth transmitting to students.

In an earlier day, the phenomena of natural philosophy—mechanics, heat, sound, light, electricity, magnetism, astronomy—all seemed distinct.

Physics now collects them in a few concepts and mathematical relationships. At one time, the sun, moon, and stars were thought to be made of some ineffable stuff, unlike anything on earth. Galileo's first observation of the moon through his telescope and Newton's demonstration that the matter of which it is made behaves like the familiar matter of the earth were the precursors of visits made with confidence that man would neither be eaten by dragons nor vanish in a puff of gamma rays. The extension of earthly physics out to the most remote galaxies scientists have detected and understanding them from a laboratory knowledge of atoms, nuclei, and subnuclear particles poses problems but finds no contradiction. Today, the phenomena of chemistry and biology are described with greatest depth and clarity in the terms and by means of the instruments of physics, although their more complex aspects, especially those of biology, require also the specialized concepts and techniques of the particular discipline. The practical arts, which depend heavily on familiarity with materials, grow increasingly less dependent on empirical knowledge and more dependent on theoretical understanding gained through physics. It seems fair to say that one of the most far-reaching consequences of the pursuit of physics has been a continual unification of knowledge, which makes it possible for man to retain the concept of the wholeness of nature while exploring it in all its rich detail. This consequence must be counted a humanistic achievement of high order.

Direct experience with phenomena and measurement is important in learning physics. In an introductory physics laboratory one can gain experience with a simple system, making attempts to control it, modifying it, making observations and measurements. One can muddle, err, cope, retrace, work through apparent contradictions, misinterpret, and grope toward comprehension. Assured that this devious path is the norm in science, one learns to have confidence in his ability to devise his own path and to recognize incorrectness and correctness without having to be told by someone else. Confidence in one's ability to work independently and in one's own judgment are qualities that can be nurtured effectively in a carefully planned and staffed physics laboratory.

5.2 COLLEGE PHYSICS FOR THE GENERAL STUDENT

Why is it that physicists consider their subject an exciting and rewarding intellectual enterprise, worthy of pursuit with fervor, whereas the world abounds with highly intelligent, well-educated persons who find physics dull, mysterious, or threatening? At least two attributes of the traditional teaching of physics contribute significantly to the problem: Physicists cater, in teaching, to their image of physics, which is pure and rigorous and is reinforced through their contacts with colleagues and their success with those few students who tend to see physics the way they do; for all but a few students, the pace is too rapid, and the instructors's desire to reach the modern physics at the end of the course, too great.[37]

The general liberal arts student needs the experience of learning what it means to really know something in a quantitative as well as a qualitative way. But he suffers more than the others from the inappropriateness of a long list of typical end-of-chapter problems or of lengthy, contrived laboratory sessions unrelated to recognizable experience, past, present, or future—two instructional crutches that do nothing to illuminate the genuine source of satisfaction for practicing physicists. Of all the students we are considering, those who are uncommitted to science suffer most painfully from the fast pace of much traditional physics teaching. The frustration and confusion that result from a well-meaning effort to present too broad a spectrum of physics in an introductory course must surely contribute to the suspicion, hostility, and fear with which so many students seem to view physics. A student comment, frequently quoted, is apt: "Taking physics is like trying to take a drink from a fire hose."

These problems have many obvious parallels with the teacher education problem, and their solutions are compatible. As is the case with future teachers, the general student especially needs time to sense the significance of physical law as a summary of experience to date, subject to evolving shifts in semantics and in extent of generality. These students benefit particularly from a realization of the provisional nature of scientific knowledge and from demonstrations of predictable and verifiable limits of validity, a facet of knowledge not always so readily shown in other fields.

Society faces environmental and social problems, many of which will be alleviated through the application of fundamental knowledge that is already known or not yet available. In the political arena, there is too little understanding of the structure on which technology, including the technology needed for control purposes, rests. One often wishes that more public servants could have had some contact with physics at the college level, contact that would have included an opportunity to gain insight into this structure. The so-called informed public also should be better educated in this subject. It is tempting, therefore, to orient courses for nonscience students toward current problems of major concern. This approach, however, risks a superficial sampling of the basic material. It is more appropriate to teach within the guidelines suggested earlier, providing some ability to understand natural phenomena: for example, why are there aurorae; why do Polaroids help one see fish in water; why are sunsets red; why are no stars visible in the daytime; why does stopping a car traveling

60 miles per hour require four times the distance needed to stop one going 30 miles per hour; how can a sailboat go upwind; or even why is there thermal pollution from power plants? It is also important to help these students understand the motivations and social structure of science as well as recognize that many types of order-of-magnitude calculation are not beyond them.

The apparent affinities between physics and the humanities have been the subject of much thought and writing. Changing modes in literature, music, and the fine arts during man's history frequently have coincided with the introduction into physics of radical new modes of thought. In the case of music and art, physics offers techniques and explanations to enrich the understanding of those who wish to practice or appreciate them. Physicists often feel deep personal affinity with the arts. Such affinities can be exploited in joint courses (for example, physics and literature, physics and music, physics and art), with the potential to attract students who ordinarily would never enroll in a "hard" science. In these courses, dialogue is best assured if instructors from the two disciplines are present simultaneously throughout the term. Although such courses cannot offer much more than a flavor of physics, they can give students a new perspective by exposing the conscious and unconscious use of metaphor deriving from it.

The history of science has become a recognized field, and many historians of science have strong backgrounds in physics. Nevertheless, a course conducted jointly by a physicist and a historian offers the possibilities of cultivating awareness of the thinking of some of the world's most creative persons and of an enriched exploration of the influence of physical ideas on man's intellectual development. The role of physics and of physicists in affairs of the recent past offers an additional basis for approaching students who have a diversity of interests among the social sciences.

Physics faculty members often express the wish that students would study physics much as they do history or English, as a component of a liberal education. Students in the courses offered for nonscientists are the best existing approximation of this ideal. Because some of them, at least, have chosen freely, we must give their interests and needs the same quality of attention we give to those of our majors.

5.3 STUDENTS OF OTHER SCIENCES

Students majoring in sciences other than physics are entering disciplines that have been influenced by physics in striking ways. Through its own nature as the most quantitative, structured science, physics serves as a model. Through its explications of atomic structure and interactions between atoms, it provides the theoretical basis for chemistry and those disciplines that grow from or depend on chemistry. Several recognized scientific disciplines of today, e.g., geophysics, atmospheric sciences, and health physics, can be considered in many respects to be applied physics. Finally and by no means least importantly, most measurement in science, medicine, and technology today is measurement of physical properties of matter and depends either on adaptation of techniques and apparatus originating in physics research laboratories or on electronic components, transducers, circuitry, and other devices that originated from the demands of physical research. The techniques of ion kinetic energy spectroscopy, photoelectron spectroscopy, pion therapy, and semiautomatic scanning and pattern recognition, to mention only a few, are today undergoing adaptation to particular needs of other fields. In sum, there are a number of reasons for including physics in the education of scientists of all kinds.

The students in these various disciplines are guided and constrained by requirements and recommendations established by their particular faculties. These faculties confront the problem of fulfilling the practical need to provide their students with a sound and comprehensive background and, at the same time, responding to the free choice of electives. In addition, they must weigh the willingness of the physics department to offer courses that match the backgrounds and needs of their students. Rarely do physics departments take the initiative in consulting with other science faculties to gain insight into other points of view on and possible applications of their physics courses; they might well do so periodically, recalling that possibly a majority of the students they teach are directed to them because of certain expectations.

For biology-oriented students, the question of training in physics has considerable current significance. The words of biologist Dana L. Abell, written when he was Associate Director of the Commission on Undergraduate Education in the Biological Sciences, are particularly compelling. Dr. Abell stated[38]:

> An important inconsistency in planning curricula for undergraduates in biology, agriculture and natural resources has always been that training in physics is required but almost never used. The conviction that physics is an essential part of these programs is firm, however, even to the point that some people are already debating the content of a "second course," i.e., physics beyond the introductory year. These people would seem to be maintaining that physics isn't used because the students with only one year of work aren't sophisticated enough to understand the material through which an interdisciplinary exchange between biology and physics has been accomplished.
>
> CEANAR [Commission on Education in Agriculture and Natural Resources] and CUEBS [Committee on Undergraduate Education in the Biological Sciences] share the conviction that introductory physics courses belong in most of the major programs

with which they are concerned, but we are convinced that elementary physical concepts can be introduced at many points in these programs. As we see it, interfaces do exist between these fields at all levels, but opportunities for interchange short of the highly sophisticated interdisciplinary field of biophysics or in more specialized fields such as agricultural engineering have simply not been created.

Dr. Abell also made the following points that suggest that physics faculties should question how well they are meeting the needs of these rapidly developing fields:

While faculties of physics and the biological sciences both decry the watering down of this non-major course, circumstances conspire to make it a consistently less rigorous course than the one for majors. Yet, it is made to seem excessively formal even without prospective physicists in it, it is designed to fit the tastes of physicists alone.

Undoubtedly physics *is* best taught as a subject in its own right and the instructor *should* be allowed to choose illustrative material with which he will feel comfortable, but physical concepts do pervade all of science, and perfectly good examples of physical principles can be drawn from almost any aspect of biology, both theoretical and applied. The fact that this is not done relates more to the training of physics faculties and to traditions in the teaching of physics than it does to any real boundaries between physics and the biological sciences.

Changes in the teaching of all sciences are weakening traditions and lowering the barriers between fields, even at introductory levels. Information on the significance of elementary physical principles to fields outside of traditional physics which is made both accessible and appealing to the physics instructor could easily be incorporated into these non-major courses.

Greater use of illustrative materials from outside the traditional bounds of physics could have the effect not only of building a substantial bridge to those other fields, but also of raising the level of sophistication as increased student interest brings greater incentive and commitment.

Students of the physical sciences, other than physics, do not encounter problems of the kind faced by the biology oriented. Often, and with increasing frequency, students of other physical sciences find themselves in an introductory course offered for physics majors and engineering students. Few would argue that this arrangement is inappropriate. Beyond the introductory year, however, problems arise because subsequent courses, which might be suitable for other physical scientists, often are directed more to the supposed needs of the physics major than is necessary. At the same time, the faculties of the other physical sciences tend to emphasize their own professional outlooks by offering courses that cover essentially the same ground, but from the idiosyncratic viewpoints of their various disciplines. The validity of the contention that physics is basic to other physical sciences argues for intermediate physics courses congenial to those students of these other fields who wish to explore the foundations of their major subject.

5.4 PREMEDICAL STUDENTS

The long-standing custom that premedical students complete a year of physics might be a residue of ancient modes of thought and education, for the nouns "physician" and "physicist" have a common root in the Greek word for "nature." Until a century or so ago, physicians were, in fact, among the most learned men in science generally and contributed importantly to fundamental advance in physics (for example, Mayer and Helmholtz). One effect of the subsequent rapid differentiation between scientists and medical practitioners was the redefinition of the role of physics in medical studies. Ostensibly, it has been supposed to aid in inculcating a small amount of general scientific background and to induce a degree of familiarity with topics (e.g., optics, acoustics, and fluid mechanics) that could be of direct professional value. But premedical students, and occasionally members of the physics faculty, commonly suspect that introductory physics is used as a screen to sort out aspirants having a certain degree and kind of academic capability. The author of a textbook on physics for medical students wrote: "Speaking of medical physics examiners; it is deputed to us to hold an outer gate to your profession."[39]

Surely, if this were the reason for the physics requirement, we would not concern ourselves over premedical students. The earlier, more cogent reasons must be reinterpreted in the light of the changing modes of medical education today. No longer will there be a single pathway into or through medical school, for it is becoming widely recognized by medical educators that the variety of forms of medical practice today requires greater variety in student background, interest, and talent than has been customary. Routes will be provided for some students whose undergraduate preparation has concentrated on the social sciences; many of them may need "remedial" work in the natural sciences. At the same time, medical schools are also encouraging stronger preparation in the basic sciences and mathematics. This latter tendency is revealed by recent data showing that half of the nation's premedical students enroll in the physics course intended for majors or for both majors and nonmajors.[40]

Apparently, the physics requirement will continue to stand for some time as a kind of paradox. Still deemed important, as well it should be, given the strong strand in biology that leads in the direction of reductionism and the constantly increasing sophistication of medical instrumentation, the physics course is nevertheless only a small segment in the long road to either medical practice or medical research. Only about 1 percent of medical students have undergraduate majors in a physical science other than chemistry, and few medical students take more than the minimal requirement in physics. Faced with major problems that promise to alter medical practice and its organization in fundamental ways, it seems unlikely that medical educators will attempt to define their interest in a physics requirement more clearly than at present, if, indeed, they ever could.

In the meantime, physics faculties can recognize aspects of the course for nonscience majors that future physicians should encounter. Certainly, as urged by Abell, they can identify topics (e.g., radiation, waves and pulses, fluid flow, feedback systems) that promise to be significant in future biology and can add interest to lectures and laboratories.

5.5 ENGINEERING STUDENTS

Physics traditionally has been an important part of the curriculum for engineers, and necessarily so. The distinction between engineering and applied physics becomes more and more an artificial one. It is an experience shared by many instructors that, in some upperclass courses designed primarily for physics majors, engineers have been among the best (and worst) of the students. These remarks are not intended to imply that physics is a sufficient basis for engineering, but there is a strong coupling. In fact, engineering faculties have evolved in the past 10–15 years into applied science faculties, frequently staffed in part by physicists and capable of teaching aspects of physics they feel their students need. They have grown all the more anxious to do so, as physics teachers during the same period have increasingly relied on abstract formulations of physical law at the expense of concrete illustration.

The situation is further complicated by the proliferation of engineering specialties (aerospace, nuclear, chemical, civil, communication, electrical, geological, mechanical, sanitary—all these appear as departments listed in the catalogue of a moderate-sized engineering school, and each may have its own attitude toward a physics requirement).

The important distinction between engineers and physicists, a distinction that must be kept in mind in assessing their special needs, is that of outlook. The engineering tradition is more pragmatic, and less delighted with order-of-magnitude reasoning and with looking at various ways in which a problem can be viewed and solved than with learning the most expedient way of solving. For engineering students, physics is useful insofar as it provides the foundation in fundamentals that allows a systematic approach to be used intelligently. In this sense, the ability of physics to demonstrate that many complexities derive from great principles is important.

Specifically, in physics courses designed for engineers, the conservation laws should be emphasized, as should fields from static and moving sources, velocity fields, and fundamental wave theory. Engineers can understand how things work without much physics, but the reinforcement stemming from the knowledge of *why* things work comes from physics and represents an investment in the future.

Finally, the interrelationships between physics and engineering will be fostered if the courses shared by students in the two fields begin to recapture some of the flavor of application. The once-popular engineering physics baccalaureate degree lost its appeal, possibly because it left no time for the leavening influence of the humanities. If this deficiency could be repaired in a satisfactory manner, this degree might be useful in a future that will be greatly concerned with applications of science.

5.6 PARADOXOLOGY

It is always easier to condemn and criticize the establishment than it is to offer viable substitutes for present practice. The implicit and explicit criticisms contained in this chapter are hardly without precedent. Neither have they been totally ignored by physics teachers, many of whom have indeed tried many approaches, some with success. The problems, however, are truly vexing, as the following summary demonstrates. Possible improvements may be suggested but cannot be prescribed. The structure of higher education in the United States gives practically complete autonomy over curriculum to each of more than 2000 faculties. We would not have it otherwise, but this characteristic diversity guarantees a lack of agreement as to the way these complex educational problems should be resolved, even among small groups of experts focusing attention on a relatively narrow aspect of the whole picture. Local needs and strengths differ markedly, and there is much to be said for trying a variety of unrelated approaches. Nevertheless, to avoid dialogue would be either to admit defeat or to decide to leave things to chance. Fortunately, dialogue is lively in the meetings and the publications devoted to physics education.

It is well to face the most awkward issue squarely. We bring it into focus by developing some paradoxes that are implicit (and sometimes explicit) in suggestions of conscientious recognized experts attempting to offer insight into the nature of the problem of teaching physics. Of course, the most basic paradox is the one that opens our discussion of college physics for the general student. Proceeding from this one, what is more reasonable than the following:

Our courses are presented at much too rapid a pace to allow for the assimilation of ideas, for how concepts and theories originate and are validated, much less for allowing reflections on the scope and limitations of scientific knowledge or its impact on our intellectual heritage and the development of a world view. We must cut down on coverage.

But:

Current introductory courses must contain a larger fraction of "modern" physics than heretofore and must become less abstract and more applied than before. Early in the course, we should present special relativity and qualitative quantum concepts that "turn students on."

But:

Students are being told about the "fascinating" particles of high-energy physics with jargon about interactions, angular momentum, mass–energy, uncertainty principle, while they have not yet achieved an understanding of what is meant by such things as velocity, acceleration, force, and charge.

But:

Why are we professionals excited about physics? It is partly because of the insight it gives us into the structure of the physical universe. We are no longer fascinated by bodies falling from towers or by calculating currents in circuits.

But:

Students must master the fundamentals of mechanics before being taught planetary motion and should know optics before being taught about lasers.

But:

It is not true that if you teach only the basic principles to students they will see how to apply them to real problems.

But:

Students can acquire intellectual perspectives toward the methods and concepts of physics, but these are not automatically conveyed by training them to calculate how high a stone will rise when propelled into the air.

On the other hand:

There are many ways in which modern man must interact with a highly technical society. It is difficult to understand the evening news report without some understanding of current technology. Informed citizens must be able to know more about technical matters to evaluate questions of national defense. In addition, just the ability to understand better the various natural phenomena one sees in life is a major contribution of a course in physics.

But:

> There are also, of course, the topical pressing social problems that stem from the release of nuclear energy, the application of science to warfare, and the problems of controlling the abuse of our environment. The significance of these vital problems should not be minimized, but an adequate foundation must be laid before intelligent, critical discussion can occur; otherwise it is no more than trivial, superficial chitchat.

What are the identifiable points in these nearly universal, albeit contradictory, criticisms of the usual physics courses, and what suggestions for improvement can survive them?

First, it is clear that the traditional concept attempts too much too fast. "Less may be more" has been quoted as a proper slogan for introductory courses.[41]

Second, there is too much emphasis on verbal presentation and mathematical formalism. "Curriculum revision" too frequently has consisted of moving material down from the graduate to the undergraduate level. Such revision misses the needs of the students considered in this chapter. Appeal to phenomena through conceptually transparent lecture demonstrations and laboratory experiments is more to the point.

The third point is the insufficient emphasis on interesting applications of physics. The undergraduate program is an appropriate place for a shift in emphasis.

In short, we must recognize that we can no longer afford to project our own pure image of physics at all of our students, regardless of their needs and interests. Few of today's physics majors are ready for the cold, clear light, and even fewer of the others are prepared for it.

6 Technical Physics

The reasons for the phenomenal growth of the community colleges (from 680 colleges and 660,000 students in 1961 to 1100 colleges and 2,500,000 students in 1971) are many and varied: open admission policies, favorable geographic distribution, low tuition and fee schedules, and varied program offerings appealing to a wide variety of post-high-school-age groups in the population. For our report, their growing importance as the home of technician education has special significance.

The community college has always been—at least in theory—an open-door college and a microcosm of the community it serves. Hence, it is not surprising that in its population we find students more representative of the entire college-age population of the United States than are those in any other segment of higher education. Community colleges located in or near the inner city of large metropolitan areas often have a substantial percentage of entering students with below-average ability as measured by high school performance. Substantial fractions of these students are from minority groups; in fact, practically all of the increase (in percentage terms) of black enrollments in higher education in the last couple of years has been at the community college level.

Owing to the multiplicity of missions to which the community college is dedicated, there is an obvious need for a variety of different physics courses. Ideally, each of the objectives of the community college (transfer, technical, remedial, general, and adult education) demands a different approach and a separate course. Except in the largest community colleges, diversification on this scale is not possible.

While some of these objectives are held in common with senior colleges and have been discussed frequently, the peculiar needs of the technician have only recently been perceived at all clearly by physicists in the United States.[42]*† It is not surprising that little of the teaching in physics has been responsive to these needs. All too frequently in the community colleges, the traditional introductory college physics course has served as a model, albeit dressed up (or dressed down) for consumption at the two-year college level by the addition of special supporting materials, such as problem books, teacher's guides, and laboratory manuals. The mismatch between expectations and reality is evident and often produces unfortunate results.

6.1 PHYSICS FOR THE TECHNICIAN

Recently, the Bureau of Labor Statistics issued a report projecting technical manpower needs to 1980. It states that "Physics technicians and Mathematics technicians are expected to show the fastest growth rates among the technician occupational specialties, 95 per cent and 91 per cent, re-

* It will be of interest to physicists that Kamerlingh-Onnes initiated a nationwide program for training physics technicians in The Netherlands at about the time he commenced his research into low-temperature physics.

† The October 1971 issue of *Science Education News* (American Association for the Advancement of Science) is devoted entirely to technical education.

spectively."[43] The projected demand is for 20,700 "physics technicians" by 1980, an increase of about 100 percent over the number in 1966. For the more inclusive category "Engineering and Physical Science technicians," the need is projected to exceed 200,000. Bureau of Labor Statistics projections tend to be controversial, and it is not our intention to comment on the degree of confidence to be accorded the report. The qualitative assumptions underlying the projections seem soundly related to well-established industrial trends: increasing utilization of technicians relative to total employment, due to the expansion of research and development activities, increasing complexity of industrial processes, and the growth of industries employing large numbers of technicians. Although recent cutbacks in industrial research may raise doubts about the soundness of these assumptions, the employment of technicians nevertheless will probably reach and maintain a substantially higher level than the present one.

In 1965, the largest source of new technicians was the upgrading of existing employees; the next largest source was post-secondary-school training of a kind offered by technical institutes, junior and community colleges, vocational-technical schools, and the extension divisions of engineering schools. The existence of many new two-year colleges guarantees that they will make an appreciably larger contribution to the technician work force than was the case in 1965.

A report by the NSF to Congress states[44]:

> The modern engineering technician occupies a position between the engineer and the skilled worker. His job is to translate the ideas of the engineer into working plans to be followed by the shopman in producing a product or carrying out a testing procedure. He must be acquainted with the associated engineering field and also with the detailed work procedures involved.
>
> The engineering technician curriculum is post-secondary, is most generally terminal and provides instruction in theory and applications related to science and technology.

The physics component of the curriculum for technicians still lacks full definition. There is a need for more suitable materials and for instructors who are well grounded in physics. These instructors must have a good understanding of the nature of the students who enter the program and of the kinds of responsibility these students will have when they enter employment.

According to the report of the Panel on Physics in the Two-Year Colleges,[42] the student technician group includes both those with college-preparatory and those without college-preparatory high school training. The typical entering student has a mathematics proficiency level below that of trigonometry and sometimes below algebra. He rarely has studied physics in high school. He is not likely to challenge a concept that he does not understand. He lacks self-confidence. Student technicians generally come from lower socioeconomic levels and find less family and peer group support for their studies than do college students. Tests indicate a wide gap between the quantitative and the verbal abilities of the technical student, the verbal being much the lower. According to the report, one technician made the following comment, which characterizes the aspirations of many potential technicians:

> I found that my high school diploma did not qualify me for a good job, and I thought that two years at . . . would qualify me for a better job. I chose the technician program rather than the liberal arts because I enjoyed fooling around with my hands, building things, rather than reading books.

It should be noted that such strong motivation can produce striking results.

There is a divergence of opinion that intrudes on discussions of physics education: Should the emphasis be given to basic concepts, on the assumption that, when understood, their application can be left to the student (or in this case, learned on the job), or is it necessary to stress application in order to teach physics effectively? Technician education asks that both be included, as can be seen in this portion of a report issued by the U.S. Office of Education[45]:

> The technician must have sufficient knowledge of the basic principles and phenomena of the science underlying his specialty to be an effective, comprehending, and perceptive worker with his or her professional counterpart and to be able to master the inevitable (and often rapid) changes brought about by technological developments.
>
> It should be assumed that the professionals in the field supply the deep theoretical components of the task, but the technician must be sufficiently grounded in the fundamental principles to permit some interpretation of phenomena he encounters, to have a sound understanding of the theory as it is applied in the field, and to learn of technological changes in his specialty by independent study of reports of developments as they occur. The basic science courses in his curriculum must provide the knowledge of the scientific principles and their application needed by the technician.
>
> The basic science courses for the physical science and related engineering technologies are fundamental and applied physics, and usually some study of chemistry; these form the base for specialized courses in mechanics, statics, strength of materials, electronic circuitry, instruments and measurements, and other specialized applied physics as required by the particular technology.

Because of the characteristics of the student in a technical curriculum and his expectations from physics, the usual introductory college courses are unsuitable. Most technical physics instructors agree that there is a serious dearth of useful instructional material. They agree, also, that their courses need to focus initially on things rather than on abstractions: machinery, equipment, and instruments with which technician students already are acquainted or are to become acquainted. The mathematics they study should be tied closely to their physics and other technical studies. Laboratory work should be emphasized, with a written report as an important element of each exercise. Problem solving, involving practical problems related to industry and engineering, also needs emphasis. The move toward generalization and abstraction should be deferred, although it must be made. This progression of emphasis is similar to that advocated for essentially all physics students, although for the technician student it needs to be honored more strictly.

The desirability of breaking away from the traditional time quanta has appeared again in discussions of this curriculum. Here it is explicitly manifest in a proposal to offer laboratory work in modules involving from one to three weeks of instruction centered around a physical system of some sort, for example, a carburetor, or an engine test bed. A careful distinction should be made between the use of such a system in a physics course and a technology course. In physics, one would go from the specific example to a general principle and then back to the example, seeking the manifestation of yet another general principle. In the technical course, the system is the principal object of attention. Sample modules have been prepared* and are being tested.

As is the case with physics curricula more generally, the objectives of technician curricula vary from institution to institution, and in any given location there will be a spread of student background and ability. No single prescription can meet all of the needs, so those who gather to discuss and develop technical physics courses are obliged to prepare materials that can be adapted to varied circumstances. In the two-year colleges, where most of the training of technicians will be accomplished, the problem becomes intensified, for physics instructors must also offer courses that prepare students for transfer into four-year colleges. The often inadequate working conditions and the heavy teaching assignments in two-year colleges exacerbate their problems further.

One outcome of the 1969 meeting[42] was the establishment of a National Steering Committee to stimulate and coordinate the production of physics instructional material for technicians. The committee was initially supported by a grant made by the Esso Foundation to the American Institute of Physics. The American Association of Physics Teachers has created a Committee on Two-Year Colleges to carry on the work initiated by the Commission on College Physics. The interests of this Committee go beyond but certainly include the improvement of technician education.

* Production centers, supported by the National Science Foundation, are located at The Technical Education Center, Cambridge, Massachusetts; Florissant Valley Community College, St. Louis, Missouri; State University of New York at Binghamton, New York; and Oak Ridge Associated Universities, Oak Ridge, Tennessee.

6.2 THE TWO-YEAR COLLEGE TEACHER

According to the American Association of Junior Colleges, about 60 percent of community college physics teachers taught only physics courses.[46] Thus, a substantial minority of community college physics teachers also find themselves teaching such courses as chemistry, mathematics, and earth science, especially at the smaller institutions.

Many of these "part-time" physics instructors are primarily teachers of chemistry, mathematics, and the like and teach physics only because it is required of them. This is borne out by the 1967 NSF study, *Junior College Teachers of Science, Engineering and Technology,*[47] which found that over 35 percent of all community college physics courses were conducted by instructors with twenty or fewer undergraduate credits in physics, that about 7 percent of all courses were conducted by instructors with no graduate credits in physics, and that over 40 percent of these courses were conducted by those with twenty or fewer graduate hours in physics. Many, if not most, of these teachers need to receive additional training, through participation in summer and in-service institutes planned in large measure by successful, highly qualified community college physics instructors.

The *Registry of Junior College Science and Mathematics Teachers* also notes that only about 10 percent of those community college physics teachers who teach only physics courses teach exclusively in the transfer program. Hence, a substantial majority of the community college physics teacher population teach both transfer and nontransfer level courses in physics. Since, as we have stressed in Chapter 5, no single method of instruction can be used successfully with such a heterogeneous group, the technical physics teacher also will have to be familiar with a variety of forms of educational technology. To promote experimentation, institutions training teachers for two-year colleges should give much more emphasis to learning techniques such as audiotutorial units, open laboratory experiments, and modular structure. The prospective teacher could undergo the experience of taking such a nontraditional course from a successful teacher or of serving as an intern. Degree programs designed specifically to prepare community college physics instructors are being introduced at several universities.[48]

6.3 ARTICULATION

Our concern with education in technical physics has obliged us to single out the two-year college, where almost all of formal technical education is located. It becomes appropriate, therefore, to mention another matter in which the two-year college is involved, the relation between the two-year and the four-year college. On the one hand, it is the responsibility of the four-year colleges and universities to prepare young physicists for careers in community college teaching; on the other hand, the community college must be able to prepare a substantial fraction of its heterogeneous student population for transfer to a four-year college at the junior level. Obviously, the transition from a community college to a senior college involves some problems.

It is consistent with the ideology that undergirds the community colleges that physics courses be taught as we have advocated in Chapter 5.

Even so, a difference in the level of difficulty at the two types of schools is likely to appear. Upon arrival at the university, the transfer student can find himself ill prepared. Courses bridging the gap between the programs at the two types of institution are needed; probably the university can more effectively offer these courses.

Above all, the physics personnel at both types of institutions must continu to cooperate in a discussion of the articulation problem. Each side can learn from the other. Approximately ten cooperative programs involving physics faculties of both two- and four-year institutions have been initiated throughout the country, primarily with financial support from the NSF College Science Improvement Program. A way must be found to make such interaction a normal, continuous process even without federal funding. Perhaps the university involvement in preparing instructors for the two-year college implies a budgetary responsibility for assuring that some follow-up mechanism is established.

7 The Undergraduate Physics Major

The setting of our teaching is the conventional four-year college program leading to the BA or BS degree, in any college or university, as it has developed in the United States since the early years of this century. Recently, with ever more students seeking, and finding, ways to break out of the fixed mold of a rigid undergraduate program, it has been repeatedly suggested (e.g., by the Carnegie Commission) that the four-year program should not be regarded as sacrosanct.

In the next era, which promises to be one of unprecedented educational flexibility, institutions may be moving toward offering a range of college programs of varying lengths, from three to five years, depending on the particular objectives of a given curriculum. These issues touch on such mundane aspects of the educational system as the academic calendar and the place of the summer term in it, the multiple use of the available facilities, such as classrooms and laboratories, and, more generally, the grave economic problems of colleges and universities. In the light of the suggestion that three or five rather than four years in college may better satisfy the needs of young people today, physicists should address themselves to the task of constructing and evaluating models of a variety of undergraduate programs.

A novel approach to college instruction that also deserves attention in terms of the efficient utilization of material resources and student as well as faculty time is the concept of "concentrated study." This consists of a period of several weeks devoted almost exclusively to the intensive study of one subject and is obviously well suited as a component in an academic calendar based on two four-month periods of concentrated classwork separated by a one-month period devoted to special projects. (As an example, see the report by Parlett and King on an experimental concentrated study course on oscillatory and wave motion.[49]).

More generally, there are strong indications that sharp and insistent questions about "productivity" and "fiscal accountability" will be asked of educators in colleges and universities in the next few years. While physicists have in the past not felt the need for a tight national organization of their educational efforts and have refrained—in our opinion, wisely—from establishing a system of accreditation, they have found it desirable to create, at least for puposes of information exchange, an organization specifically concerned with improvement of physics education. The Commission on College Physics served this function well during recent years, but it has now ceased operating. As we become more than ever involved in seeking answers to tough questions about the efficiency of physics education, the American Association of Physics Teachers, with its new Council on Physics in Education, has the potential to become a national focal point for the educational concerns of physicists, especially at the college level. Chapter 9 treats these organizational concerns.

As new approaches in undergraduate education are explored and effective new programs with a specialization in physics are designed, physicists must keep in mind the present level of secondary education in the United States as well as the general objectives of undergraduate education. For students specializing in physics, these objectives are most expediently defined in terms of the preparation that college is supposed to provide. In the educational pyramid, undergraduate education is expected to shape the young adult intellectually so that he is broadly equipped for the rigors of the final stage in his career training, but at the same time, in the United States, it is supposed to help him to orient himself through social and humanistic modes of experience and understanding. A physics major, completing a BA or BS degree, should have studied the main areas of physics in some depth, but he is not yet expected to be a professional physicist.

The curriculum designed to meet these objectives in the physics departments of our colleges and universities is subject to many local variations, but a general pattern is easily discernible. A common basis is first established by an introductory course covering the full range of physics, now always with calculus, and extending over a period of between one and two years. A course of this kind is usually characterized by referring to a standard textbook, which is either used in most courses of this kind or which colleges feel should be used "if only the students were good enough." Nowadays the magic names are Halliday and Resnick, but only a decade ago Sears and Zemansky filled the role of standard bearers, and no doubt other fashions will follow in the future.

In the past 15 years—following the flight of the first Sputnik—a host of innovative projects and programs have sought to infuse freshness into the traditional approach to the study of elementary physics, whose character had, in spite of many changes in substance, remained surprisingly stable since the nineteenth century. The new curricula and courses all contribute in different ways toward making college physics more interesting, more contemporary, and more challenging, and toward making the teaching of physics more effective. Largely in response to the notable improvements in secondary education in America during the same period, especially in mathematics, all of the new physics courses for science majors aim at "beefing up" the material presented at the introductory level.

Of all the efforts made on behalf of a renewal in the teaching of introductory physics, none has exerted a greater influence on the physics community than the course given by Feynman and published in three volumes as the *Lectures on Physics.*[50] Probably no book on elementary physics has ever graced so many reference shelves in physics libraries and been studied by so many professional physicists—teachers and nonteachers alike. Feynman's lectures are unquestionably the most original contribution to the textbook literature in many years, and his insights into physical processes have already inspired and influenced a whole generation of physicists.

A very different new approach to the teaching of elementary physics for science majors is embodied in *The Berkeley Physics Course,*[51] developed in the 1960's. This course comprises five nearly independent sections (mechanics, electricity and magnetism, waves and oscillations, quantum physics, and statistical physics, each written by a different author or group) and departs from the tradition of teaching all the diverse areas of physics in one comprehensive course, often presented by a single lecturer. The Feynman lectures and the Berkeley course have served as models for other intensive introductory physics courses, but because their masterful display of virtuosity intimidates too many students, neither has often been assigned in introductory courses at the freshman and sophomore levels.

It is certain that the pedagogy of physics has benefited enormously from what may be termed the "new physics" movement reflected in the preparation of the various new courses. But it seems likely to us that the next few years will not call primarily for the design of ever more exacting introductory courses. Rather, there is need to separate the good from the not-so-good by the test of experience, to deploy the worthwhile innovations throughout the educational system, and to plan carefully for the curricular needs of those physics students who are now eliminated from competition by the exacting demands typical of the new programs. The beginning of a trend in this direction can be seen in the recent publication by reader-sensitive commercial publishers of shortened and less ambitious versions of some leading introductory textbooks. In a similar vein, it is suggested by the designers of some physics courses for nonscience majors that these may well serve also for a substantial portion of the science majors who find the pace too demanding in the high-pressure physics courses.

7.1 CANONICAL SEQUENCE

Building on the foundation laid by the introductory physics sequence, the dominant curricular trend has been to follow up in the junior and senior year with a more or less canonical sequence of courses that deal in greater depth with the special areas of physics. Thus, most of the topics are taken up again, but this time with the use of differential equations, computer programs, and some reference to current research. Such a sequence of courses typically includes classical mechanics, electricity and magnetism, optics, thermal physics, electronics, and quantum physics. The main variations between departments concern the length of a particular course (e.g., the choice between one and two semesters of classical physics) and the amount of "modern physics" and its various branches that is injected.

It may be worth noting that very little effort has been devoted in the past few years to modernization and experimentation at the upper undergraduate level in the education of physics majors. Few schools have succeeded in establishing alternate routes through the junior and senior years toward the degree in physics. Since the standard curriculum has proved to be quite effective in preparing students for graduate school, there has not been much incentive for innovation at this level. As physicists become more aware of the importance of opening the major to career directions other than the conventional ones leading through the PhD in physics, it will become desirable to scrutinize the offerings at the advanced undergraduate level. But even students bound for graduate school may profit from some rethinking of the traditional patterns of the junior–senior course structure—not because tradition is necessarily suspect, but because it here embodies too rigid a subdivision of the field and thereby tends to obscure the unity of physics, which should remain a central theme in all physics education. A good example of some recent progress in this direction is the appearance of thermal physics as a standard course, reflected in the publication of several fine textbooks, to take the place of separate courses in thermodynamics, kinetic theory, and statistical mechanics, and coupling the subject strongly with the physics of condensed matter. The time may well be ripe for a review of other conventional courses, such as mechanics (should and could classical and quantal mechanics be combined?), electricity and magnetism (should there be more account taken of plasma physics?), and optics (the renaissance of which is only beginning to be acknowledged in our teaching).

7.2 THE PLACE OF QUANTUM PHYSICS

An important unresolved question facing physics departments in their educational work concerns the respective roles that classical and quantum physics ought to play in the undergraduate curriculum. It is customary to insist that a good foundation in classical physics must be laid during the sophomore and junior years, and on this foundation is placed a "frosting" of "modern" physics, capping the undergraduate curriculum with a course or two in quantum physics, atomic physics, and perhaps nuclear and solid state physics. Above and beyond the force of tradition, there is much to be said for this curricular structure: Physical measurements are made mostly with tools that obey classical physics; the physical intuition of young persons is developed by observation of the world surrounding us, which is largely governed by the laws of classical physics; and many of the technological applications for which the study of physics prepares the students deal exclusively with classical systems.

On the other hand, there are serious arguments to be made in favor of bringing students as early as possible in contact with the concepts of quantum physics. It is difficult to believe that anyone sufficiently interested in physics to declare himself a major will not be aware of the quantum notion either through high school study or through his own reading. A large fraction of current work in physics is inaccessible to anyone who has no background in quantum physics; the interaction between physics and many neighboring disciplines (notably chemistry) is dominated by quantum ideas and atomic concepts; and the most significant contributions made by physics to man's entire philosophical outlook during the past seventy years have grown out of quantum physics. Textbooks treating quantum physics at the sophomore level in science or engineering are becoming available.

It is not surprising that the physics curriculum at the college level has not yet fully adjusted to the impact of quantum physics. It took many years after Newtonian mechanics was established to develop a "modern" curriculum in which the "new" mechanics occupied its proper place. Imperfect as this analogy is, it is valid in at least one respect: The Newtonian revolution, like the quantum revolution, had a strong theoretical and mathematical flavor. As a result, the education based on it tended to be formal and abstract. Similarly, the great success of quantum physics has placed extraordinary, and perhaps excessive, emphasis on theory and formalism, not infrequently to the detriment of the development of physical intuition and structural understanding.

7.3 BACK TO THE PHENOMENA

The questions raised above are indicative of some even deeper concerns with the general style of physics teaching at the intermediate (and also at the introductory) level. Obviously, the most striking characteristic of the physics curriculum is its hierarchical structure, in which linked chains of prerequisites guide the student from the bottom to the top, with even minor detours taken only at the student's peril. The emphasis in this strong and well-tested educational edifice of neatly arranged sequences of courses is on the rapid acquisition of tightly organized information, with a premium placed on the economy of thought made possible by the use of mathematics. Memorization of facts, never essential to the learning of physics, has all but disappeared, and our present-day curriculum has a strong theoretical and mathematical flavor.

Not surprisingly, the increasing tendency of undergraduate physics education toward emphasis on and reward of manipulative skill in handling the formalism of physics has brought forth loud calls for a "return to nature." It is said that in the teaching and learning of physics, before tackling physical problems with mathematical techniques, greater attention must be paid to intelligent observation of natural phenomena, to the achievement of interpretative understanding, and to the acquisition of a modicum of physical insight. Indeed, the charge has been made that the prevailing physics curriculum has the effect of turning students into mindless equation solvers instead of developing their intuitive faculties as observers and interpreters of the world around us.

We have already referred to one reason for this trend—the nature of quantum mechanics and its impact on most of physics—but there are other causes as well. It may well be that the preoccupation of many physics students, and some physics teachers, with the purely formal aspects of physics mirrors the process of maturation of young adults. To look at a problem from the proverbial, and somewhat subjectively defined, "physical" point of view may seem like the natural, innate, human approach. Actually, it requires time, patience, and experience to develop a physical intuition, and while good teaching can help a student substantially in this process of maturing, there is no reason to discard the learning crutch provided by easy familiarity with formal manipulations. Excessive mathematization of the curriculum for its own sake is certainly undesirable, but the proper use of formalism to enhance rather than retard the physical understanding and clarify the physics is obviously to be encouraged.

On the other hand, there is every reason to advocate closer contact with the phenomena at all levels of physics education. This brings us to the role of the laboratory in the physics curriculum. Dissatisfaction in this area seems to be endemic. The variety of attitudes taken on this question throughout the country indicates not so much a commendable diversity as a basic state of confusion and indecision. The questions that are being raised are familiar: Should the lab serve primarily as an instructional device supporting and illustrating the lectures and recitation, or should it be fundamentally regarded as a series of minute research projects, confronting the student with an opportunity to sharpen his manual skills and to show him what doing "real" physics is like?

In the first case, there is a premium on having a great variety of experiments, many of a rather qualitative nature, illustrating the more important principles introduced in the course, and the equipment is best ready-made so that a minimum of time is spent by the student on design and construction. In the second case, it is far more important to have a few experiments that are designed to provide accurate quantitative results without making excessive demands on the student's time. There are, of course, laboratory programs in operation that combine these two extreme viewpoints, and a spectrum ranging from the conventional course-centered lab that parallels the lectures to the open-ended lab and the laboratory in the department's corridor.*

Many of the problems posed by the teaching laboratory are economic. Laboratory instruction is expensive not only because of the use of equipment and supplies but even more so because of the requirements of close

* See, for example, the reports "The Divergent Laboratory" by J. W. G. Ivany and M. R. Parlett, and "The Instrumented Laboratory" by J. A. Soules, bound together and distributed by the Commission on College Physics, August 1968. See also Shonle.[52]

personal staff supervision. This means heavy teaching loads in small colleges and heavy utilization of graduate teaching assistants in large universities. Often in recent years laboratories accompanying the large introductory courses have been curtailed or entirely dropped for reasons of economy compounded by general dissatisfaction, although in most cases educational rationales were found to justify the change. A rather significant modification that seems to be gaining popularity is the separation of laboratories from the junior-senior lecture courses and the establishment of separate intermediate and advanced laboratory courses. Such an arrangement removes many constraints imposed on the nature and quality of the laboratory by a close coupling to a particular lecture course and allows for more efficient utilization of available resources. In our opinion, the divorce of the laboratory from the lecture courses can be a healthy development if it frankly acknowledges that most experiments pertain to several diverse subjects in physics and if it does not lead to an undermining of the phenomenological content of the lectures. However, the separation must be consciously viewed by students and faculty as a measure intended to strengthen the laboratory component of the physics curriculum and not—as may too easily happen—as a relegation of the laboratory to secondary status, perhaps even preparatory to dropping it altogether.

In examining the merits and drawbacks of laboratory courses, one should be aware of alternative, and possibly complementary, ways of achieving the prime objective of such instruction, namely, to get close to the phenomena. In dealing with the role of physics experiments in the undergraduate curriculum, one might again, as in Chapter 3, think in terms of three distinct levels of activity characterized, according to the magnitude of effort, as 1-hour, 10-hour, and 100-hour investments of time, each answering different educational needs of the student. At the 1-hour level, the student would, perhaps rather passively, witness an experiment as a demonstration and illustration of the lecture material; at the 10-hour level, carefully planned laboratory experiments would be actively carried out by the student, following a more or less prescribed procedure; at the 100-hour level, the student would give a substantial fraction of his time for an extended period to a major experimental project, possibly involving him in a research enterprise under faculty supervision. Ideally, the total undergraduate experience of a physics major might include a blend of all of these modes of encountering observation and measurement in physics. In particular, we advocate that the use of appropriate lecture demonstrations not be confined to the introductory courses but be extended to the junior-senior level wherever and whenever possible.

Participation of undergraduate physics majors in original current research should, we believe, be strongly encouraged. There is a vast amount of interesting "little physics" that can be done at modest expense, and the possible educational returns can be dramatic. An advanced undergraduate who works on a piece of research is likely to acquire a clear notion of what his commitment to physics means. He will have far more opportunity than usual to learn from direct contact with his professor; he will share in the excitement of exploration and in the frustrations that arise when things are not going as well as they should. Above all, at an early stage, he will come to see physics as a growing science, an open-ended enterprise, and one in which not all issues and problems are settled yet. For such research experience at the undergraduate level to be genuinely profitable, it is absolutely essential that the research problems be intelligently chosen. They must be of limited scope, and although they could certainly be collaborative in nature, they should not as a rule consist of minute, disconnected portions of large research projects. One should be under no illusion that the selection of suitable research problems for seniors is an easy task, but we regard as unduly pessimistic and restrictive the frequently heard assertion that undergraduate research is wasteful of student and faculty time because "the student does not know enough to do anything worthwhile." The current period of declining enrollments in undergraduate physics programs seems to us ideally suited for the establishment of the senior research project at least as an elective component of the physics major or as an incentive for honor students.

7.4 GRADUATE SCHOOL OR ELSE?

Until quite recently, most undergraduate physics curricula were designed for one primary objective: preparation of students for graduate work in physics. The number of graduating seniors going on to graduate schools and the quality of the graduate schools to which they were going was used as an easy measure of quality of an undergraduate program. It has now become apparent that the role of undergraduate programs in physics must be much more broadly conceived. Such programs must be made attractive to students who do not plan to pursue graduate work in physics, who might find a strong physics background useful in the graduate study of other disciplines, or who might not take any graduate work at all, seeking employment instead in industry or government or elsewhere after graduating from college.

When undergraduate physics curricula were entirely oriented toward preparation for graduate school and most physics departments were striving to become ever more involved in research, there was pressure to engage in a continual escalation of academic demands placed on students. In schools where two or more alternative major programs in physics were available, the less demanding and less professional AB programs could often be seen to atrophy and even wither away, because it was no longer fashionable to aspire to anything other than the pregraduate bachelor's degree. The same attitude also began to affect entire educational institutions. In an atmosphere that placed a premium on training the heirs to science, universities with strong research activities had a natural advantage over small liberal arts colleges, and it is not surprising that this gap grew rather than diminished. A smaller fraction of physics graduate students now comes from the ranks of degree holders graduating from small colleges than was true even ten years ago.

The current period of consolidation—forced upon the physical sciences by economic conditions—may provide unforeseen opportunities for focusing attention on different educational objectives and for reversing some of the trends of the past two decades. Undergraduate physics major programs are becoming less rigid and more available to students who do not intend to become professional physicists and to those who start with less than optimal background, culturally or otherwise. In particular, it would be well if physics departments did not treat as stepchildren those who, instead of following the royal road to graduate work exclusively, prefer to take a mixed diet of courses, including perhaps some that are applied or cognate. Successful answers to this challenge are being developed in a number of physics departments.

In the light of this analysis, different institutions should not feel compelled to seek uniform aims but should instead realize that the nation needs diversity in education. And, in any given school or institution, students should have the opportunity to prepare for a variety of interests and careers.

An undergraduate degree in physics can provide a solid foundation for a great number of different professional careers and advanced study. This is easily demonstrated by perusal of any modern textbook in chemistry, biology, geophysics, or the medical sciences. It holds as well for all technical fields and engineering. Students with a good knowledge of physics will be prepared to gain a depth of understanding of the principles underlying all of these disciplines that cannot be matched otherwise.

A special opportunity arises for those disciplines that were not long ago regarded as distinct and almost independent but now are merging into a more integrated one, often called earth and planetary physics. This broad and vigorously pursued area of investigation requires an understanding of all of physics. Advances in physics and in technology result in a continually pressing demand for their application to the observational problems of the area, and require the researcher to rely continually on basic physics in order to interpret his data. With little or no distortion of the realities, it may be said that this entire field is in principle an application of physics and in fact is recognizing itself to be so. A strong indication of this tendency is its eager reception of physics baccalaureates into its graduate programs. Physics departments have good reason to respect the intellectual links between the core of physics and these important applications, paying attention to their implications for courses, seminars, the informal activities of students, and, wherever interesting and feasible, joint or exchange teaching.

But beyond these fields to which a knowledge of physics is directly relevant, an undergraduate major in physics should be regarded as a reasonable and desirable background for students with interests in any of the liberal arts, and it would seem most desirable if a fair fraction of college graduates in all professional areas had the background provided by the study of physics at the BA level.

If this is to be accomplished, physics faculties in colleges and universities must rethink their undergraduate major programs. In all of these, the central feature of the study of physics must be retained: An understanding of some of the laws of physics in their experimental and observational setting, arrived at by analytical reasoning and developed as a continuing insight into the workings of nature. Beyond this general educational objective, diversity rather than uniformity of undergraduate programs should be sought.

There is good reason to hope that some institutions will give more emphasis to the training of teachers for secondary and elementary schools, others will provide more opportunities for occupational preparation in practical areas, and yet others may prefer to specialize in historical, political, or social problems to which the study of physics provides one avenue of access.

The precise nature of the contribution of physics education that a particular school can make will frequently depend on the presence of great individual teachers. Probably no extensively funded collective educational improvement project, no matter how worthwhile and deserving of strong support, is likely to be as effective in making a good and competent and committed physicist out of an undergraduate student as one of those rare persons—not necessarily the most charismatic or the best lecturers—who possess the magic touch of deeply stimulating and truly educating their pupils.

Finally, as we advocate that physicists and the institutions in which they teach initiate critical reviews of past practices and established methods in our educational system, we urge the protection of the demonstrated capacity of many undergraduate physics programs in colleges and universities to provide students with a solid background for graduate work in physics. It is essential that modifications and improvements in physics education not be allowed to weaken the traditional and strong pregraduate preparation needed for the continued health of physics.

To this end, we favor widespread open discussion of undergraduate curricula and practices. We urge attention to the discussions on these topics that are held under the auspices of the American Association of Physics Teachers and the articles that appear in its publications. We urge that Association to publish at periodic intervals surveys of undergraduate curricula and practices, so that faculties and students, both prospective and those in residence, can inform themselves about the state of undergraduate physics education. It is not our intent to promote uniformity of content; we have stated often in this report that diversity is one of the great strengths of education in the United States. We believe, of course, that uniformly high quality is an ideal to be vigorously pursued.

8 Graduate Education in Physics

It is risky to attempt an assessment of graduate education in physics at this particular time. Graduate study is intensive professional training and is thus directly coupled to the economic prospects in the profession. Whether one speaks of a crisis or merely of readjustment, it is agreed that the employment picture for scientists is undergoing drastic changes, and these cannot help but have a profound influence on our practices in graduate education.

As a working hypothesis, it seems reasonable to assume that in the 1970's, and perhaps well into the 1980's, no more than half of those receiving the doctorate in physics will be able to find jobs in physics teaching at or above the college level and in research. Therefore, a large number of young PhD's in physics will not be able to look forward to professional careers in which "doing physics" in the sense of advancing knowledge, together with teaching and carrying out other academic duties, is the main part of their activities. It is up to those in charge of planning graduate education to consider the educational needs of this growing fraction of graduate students. It is reasonable to expect that the shift into an era in which most professional physicists are no longer found in academic institutions will have a substantial influence on the education of professional physicists.

With the new demands placed upon graduate education barely recognized, it is far too early to say what new patterns will evolve and which of the old patterns will survive the present period of change, but it is not inappropriate to summarize the present state of graduate education in physics and to attempt to identify and perhaps evaluate current trends.

8.1 CONDITIONS OF GRADUATE EDUCATION IN THE 1960's

The end of the 1960's found graduate education in physics in the United States in a condition of great basic strength. Measured by almost any conceivable yardstick, physics training had reached a level of remarkable effectiveness, and the "system" was producing large numbers of competent, and sometimes brilliant, professional physicists. There is no need to describe in detail here the phenomenal growth of graduate and professional education in America during the past decades. As soon as a national need for a greatly increased pool of physical scientists was perceived, the universities and the federal government responded almost instantly and within ten years succeeded in tripling the output of doctorates in physics. New graduate programs sprang up in many emerging schools, and old ones were significantly expanded. This rapid buildup was accomplished mainly by a process of replication, whereby the traditional model of the graduate department at the leading universities was copied by young physicists whose attitudes toward physics had been molded by their own alma maters. While this development has contributed to the maintenance of sound academic standards throughout the university system, it has tended to inhibit innovation and experimentation at the graduate level. Without suggesting that all graduate programs are of equally high quality, one may observe that they all appear to strive for almost identical goals, differing only in degree rather than in kind in the manner by which they attempt to achieve these goals.

8.2 ENTERING THE 1970's

The extent of uniformity among graduate programs in physics is evident from an examination of the handbook on doctoral programs in physics published by the American Institute of Physics.[53] At the admission stage, the "undergraduate preparation assumed" by various departments sets some fairly uniform minimum requirements, giving students who apply with a regular bachelor's degree in physics a decided advantage over those who enter physics after the baccalaureate. Graduate school admissions requirements, including the Graduate Record Examination, seem indeed to be quite effective in eliminating or discouraging those students who do not possess enough ability and perseverance for the tough obstacle course of graduate school.

The next decade offers an opportunity to re-examine admission criteria because physics departments are finding graduate enrollments decreasing and at the same time they will be seeking to prepare students for a far greater diversity in careers than heretofore. The present screening methods favor those students who learn well from formal courses, who are conscientious and hardworking, and who are not likely to flounder in graduate school. Although it is not easy to document the contention that truly creative individuals have been stifled in their development toward becoming successful physicists, admissions policies and the rigidly structured program of the first two years may tend to discriminate against the inventive experimenter with a keen intuition but a disinclination toward mathematical formalism and against the thoughtful theorist with a philosophical bent but a distaste for the routine of problem solving. Einstein's claim concerning the deadening effect of our conventional formal education should serve as a constant reminder of the damage that might be unwittingly inflicted.

Perhaps a more serious problem than the bias of the graduate system against slow learners and against students with experimental, rather than mathematical, talent is the fact that it has been made very difficult, if not impossible, for students in cognate areas to switch into productive careers in physics. A generation ago some of the most creative physicists came from the ranks of engineering, mathematics, or chemistry. If this is no longer true today, it is probably a result more of the enormously increased specialization than of any deliberate restrictive policies on the part of physics departments. Nevertheless, efforts should and could be made to encourage the crossing of fields; and, in any event, physics departments should welcome as graduate students talented applicants with undergraduate majors in allied subjects.

All phases of graduate education, including admissions, are seriously affected by the posture of the federal government with regard to the financial support of graduate students. Here, profound changes are currently in the making, and the familiar attitude that it is in the national interest to provide substantial and direct federal support for the education of graduate students in physics and other science is quickly losing ground. Instead, it is being argued that the individual student himself, rather than society as a whole, is the chief beneficiary of a physicist's training and the direct fellowship support for the majority of graduate students is no more appropriate than support for law or medical students. These changes are likely to alter the complexion of the graduate student body by introducing elements of motivation and inhibition that have not been felt for many years.

The effect of the changing federal economic policies on graduate training will be discussed in another chapter. The primary concern of this section is with the educational aspects of graduate training.

8.3 THE CORE

The first two years of graduate work constitute the educational backbone of the PhD program at American universities. It is in the rigorous and comprehensive course work that the American student catches up with, and often surpasses, his foreign contemporaries, who usually receive more specialized education at an earlier age.

The core courses that are central to the graduate physics program appear to have their origins in the great pedagogic tradition of theoretical physics, from Kirchhoff and Tait to Sommerfeld. Once again, this time at greater depth and learning more powerful techniques, the student ranges over the central subjects of physics. This final cycle has been of greatest importance for physics, not only because it provides a foundation upon which the student's advanced work in his specialized field can be built, but also because it has given each generation of students a common language and background, of which not only the subject matter but also the anecdotes and lore form a legitimate part. In the face of increasing specialization, the physics community should be conscious of having maintained the common core of knowledge and attitude that constitutes one of the most unique features of physics as a discipline.

The context of core courses has not remained completely static, however. At an earlier time, considerable emphasis was given to continuum mechanics, including elasticity, acoustics, and fluid dynamics, and to optics. After World War II, the emphasis shifted in order to give students the background to pursue the apparent directions of research; i.e., the "core" was perceived differently. In part to avoid overly extending the period of classroom study and in large measure because the main thought of physics no longer seemed to require this emphasis, continuum mechanics gave way to quantum mechanics and its applications and extensions. The intent of a "core" program is to ensure study of those topics that are basic and broadly powerful in application and that underlie major unsolved scientific problems. Since there were and still are such problems in the field of continuum mechanics, it is not surprising that from time to time they appear in the study of phenomena newly come to prominence or in an effort to make technological advance. We now are seeing fluid dynamics in this light, and many are calling for returning continuum mechanics to the core.

It seems most worthwhile to preserve the concept of a core course of study in order to stress the unity and common elements of all of physics. We must not be too rigid in defining the core, but we must not allow it to respond lightly to winds of fashion in physics or to immediate demands of the marketplace. Since physics faculties, with all their diversity of interests and circumstances, do indeed desire to provide a broadly powerful foundation for their students, they tend to be conservative when re-examining the core from time to time.

8.4 A CALL FOR RESTRUCTURING

The normal pattern of graduate education in physics during the past few decades has by and large embraced a single model. Two years of course work are followed by an average of three years of research leading to a PhD dissertation. The master's degree, which might be acquired after about two years of graduate-level work, has been severely diminished in prestige and has become either merely a way station in the progress toward the PhD or a consolation prize for those who are to be discouraged from going on. Students in the latter category often transfer to other schools and proceed eventually to the PhD anyway.

It is difficult to change patterns such as these that have become ingrained and involve institutional egos, but there seems little doubt that the present period of ferment in higher education calls for serious re-evaluation of entrenched attitudes and for the development of new models. Under the slogan "Less Time, More Options" the Carnegie Commission on Higher Education[54] has recommended that colleges and universities restructure their programs so that "a degree (or other form of credit) be made available to students at least every two years in their careers (and in some cases every year)." A refurbished (and possibly renamed) master's degree, based primarily on the course work of the first two years in graduate school, could produce professional physicists who might enter a career in education, industry, or government without the lengthy research experience involved in a PhD dissertation. This could provide well-trained physicists, especially for those industrial jobs that require a certain amount of on-the-job training.

If the master's degree is not the appropriate vehicle for achieving this aim, the introduction of new forms for graduate degrees should be given close attention by the physics community, and efforts to break the established pattern of the existing degree structure should be given serious consideration.

Preparation for careers in teaching at various levels should be an accepted part of graduate education in physics. More has been said about the training of teachers in Chapter 4, but it is appropriate here to applaud the emerging tendency for graduate departments to take an active interest in the preparation of graduate students for careers in teaching and in science education. It seems particularly desirable that graduate students have the opportunity to gain practical experience in teaching, not only as assistants in laboratory and recitation sections of the large elementary courses but also at the more advanced undergraduate and even the graduate levels—possibly through a team teaching approach. Formal programs designed to involve faculty in helping graduate students improve their teaching capability have been instituted by some leading departments.

8.5 A CRITIQUE OF THE RESEARCH COMPONENT

It is desirable to take a fresh look even at the educational function of research training as it has evolved in graduate schools. In thinking about this matter, one should keep in mind that much of the present strength of physics in the United States has derived from an insistence on intense specialization in graduate school. Opinions differ on how early a graduate student can and should be introduced to research—and the inhomogeneous preparation of graduate students precludes uniform rules about this—but the principle seems sound that a period of apprenticeship in doing physics (as opposed to merely learning to know it) is beneficial to all advanced students. Even a student who plans a career that involves no research is well served by spending a modest amount of time participating in on-going research in some particular area of physics. The balance between general studies and specialization in research should be a strong function of the student's ultimate goals and of his ability, but only someone who has had the experience of working for at least a short period with new problems is going to be fully aware of physics as a science in perpetual self-renewal rather than as a fixed body of knowledge. Besides, even if a person who has a reasonable degree of proficiency in a special area does not pursue his special field in his professional life, he will find the possession of expertise in some area a valuable accessory to his personal development. His perspective in one specialty gives him a permanent advantage over the total generalist who knows a little bit about a great many things.

One may accept the foregoing premise without drawing the further conclusion that every graduate student must necessarily spend a large fraction of his efforts in research. There are many career opportunities in physics where advanced study is most desirable, yet where a relatively short exposure to specialized research experience suffices and the bulk of the student's time is better spent in other modes of learning.

The possibility of changing the role of the master's degree has already been mentioned. But even the work toward the PhD, with its strong and essential research component, could perhaps be made more rewarding and effective. While the concept of apprenticeship for the research student has obvious validity, we must take care that PhD students do not spend such protracted periods of time as assistants in the professor's research group that they do indeed develop an undesirable "specialist" outlook. The educational planning envisaged here would consciously enhance rather than retard the student's progress toward becoming a self-reliant scientist. Such planning would at the same time provide the opportunity to inculcate in the PhD student attitudes that would better prepare him for survival as a physicist in the 1970's. Ways and means must be developed so that the solid research experience of the graduate student, extending over a period of perhaps two years, does not lead him to think solely in terms of a life in research and teaching, especially in his thesis field. A generally more flexible attitude toward the PhD dissertation, making it less an obstacle and more an integral part of the student's overall learning experience, could have beneficial effects both on the professional preparation of the student and on the image of the PhD, without compromising its traditionally high standards.

For example, group dissertations might be made respectable in recognition of the fact that many interesting problems are too big for a single research student and that the student–student interaction in graduate school is at least as educationally valuable as the student–teacher interaction. Or, a dissertation that calls for joint supervision by professors who work in different fields of physics might be excellent career preparation.

The importance of postdoctoral training for the young research physicist making the transition to full independence in scientific work is widely acknowledged. No criteria other than individual ability and the promise of high-quality achievement should guide the allocation of the nation's scarce resources for postdoctoral education. Postdoctoral study should be carried out in the most auspicious and productive environment, domestic or foreign, and parochial considerations should not be allowed to govern. National laboratories and even private industries should be regarded as suitable places for postdoctoral training in appropriate specialized areas.

Faced with declining federal support for postdoctoral education, universities can assume a greater share of responsibility in this area by allocating to recent PhD's instructorships with modest teaching duties, allowing time for research, the writing of papers, and the preparation of seminars. Such positions should be temporary, extending perhaps to a maximum of three years, and teaching should be a carefully monitored part of the postdoctoral training, recognizing both that one learns one's subject best by teaching it and that the students in the charge of young instructors are themselves entitled to first-rate education.

8.6 EDUCATION FOR INDUSTRY

During a brief period around 1960, more physicists were employed in industry and by the federal government than in academic institutions. By the late 1960's, the dynamic factors governing employment in the three sectors produced a reversion to the more traditional balance, but not through the traditional kind of employment, namely, as regular academic faculty. All projections of employment patterns indicate that the nonacademic sector will offer the most opportunities for physicists during the next decade.

Because the decision to enter graudate education leading to an advanced degree is a career choice, though not necessarily a lifetime commitment, the student has a right to expect that the education and the degree that he seeks can offer him a realistic preparation for his chosen career. Because that education and degree are heavily subsidized with public funds, the public also has an interest in the nature of the education being provided.

In recent years, spokesmen for industry have indicated their concern with the graduate education of scientists and engineers and, in some instances, have spoken directly about physics education.[55] * Reports of their opinions have been both published[56] and circulated informally.† We summarize them in the following paragraph.

By far the most universally endorsed theme is stated by Ascher: "I believe that the demand for the general is the most important single demand that industry can put to the University; it is also an important contribution on the part of industry to the solution of the crisis that the University undergoes."

A second theme, clearly of equal urgency and not independent of the first, is the anguish that often accompanies a mismatch between career aspirations and actual careers open to individuals. Concurrent with Ascher's statement is the view of the PhD degree as signifying predominantly training in research and in the art of specialization. This view attaches relatively little importance to the subject of the dissertation, pointing out that even if the physicist refuses to change his specialization every decade, his specialization will change. Spokesmen[55,56] † point out that industry has a strong interest in people with the flexibility to shift from one area of work to another and to careers in which they will draw on their scientific backgrounds and the knowledge and perspective gained from fields quite different from those in which they may be working.

* See, for example, Ellis[56] and A. M. Bueche, "Issues in the Changing Relationship between Industry and Academic Science" (General Electric Co., Corporate Research and Development, Schenectady, N.Y., 1970) (unpublished paper).

† E. Ascher, "A Contribution to the International Seminar on Education of Physicists for Work in Industry" (unpublished manuscript).

These general beliefs lead to more detailed opinions about the nature of graduate education. For instance, the student needs greater awareness of the role of physics in society. He may gain such awareness more through courses in which case studies are a principal part than through introductory courses in various social sciences or humanities. Students also need to know something of the style of industrial research and the complexity of most industrial problems, problems that require for their solution the integration of knowledge and techniques from a number of disciplines, often extending beyond the sciences. The long period demanded today for PhD training (6.3 years is a median in the United States for a doctorate in physical science and engineering) is of concern because it consumes a substantial interval during the highly productive years of a young scientist, thus increasing the cost to both the nation and the individual in terms of productivity.

There is a great need to strengthen bonds between industry and the academic institutions through temporary exchanges of personnel, internships, adjunct professorships, and the like. Collaboration between universities and industry could bring about successful programs for midcareer education that would keep scientists abreast of their fields, offer opportunity for those who are changing fields to acquire a new base of knowledge and skill, and give a broadened comprehension of the overall state of science and of science in its social context to those whose careers are taking them out of the laboratory.

In general, these views are not in conflict with those expressed elsewhere in our report. Industry's complaint of "narrowness" in PhD training is accepted by the academic community as directed more toward attitude than substance. Even elementary particle theorists can employ their highly developed computational and programming skills to industrial problems, bringing with them, in addition, the broad core education in physics they have received. In the United States (compared with most other nations), a relatively close coupling between industrial and academic physics is apparent in almost any research conference. During the recent past, the demand for rapid expansion of the educational system produced strains apparent to all. However, the absence of intellectual barriers assures that, in a steady-state climate, differences in view are unlikely to persist.

8.7 PURPOSE IN GRADUATE EDUCATION

In graduate education, as in undergraduate education, a call for greater diversity is being heard everywhere. There are eminently good reasons for such advocacy. Graduate schools in the United States have been expanding for almost a hundred years, and opportunities and facilities now exceed the current demand. An impulsive contraction of the system in order to bring the effort expended nationally on graduate training in line with today's demands would, in the long run, be counterproductive. Rather, American graduate education in physics should attempt to maintain its momentum by utilizing to advantage the period of consolidation that economic and political circumstances have brought about and by diversifying its resources in the interest of strengthening rather than weakening the fabric of physics.

There is a variety of ways in which our graduate education potential can be utilized more effectively. The model of an educational pyramid, suggested by H. R. Crane, with a broader base and a narrower top and offering more opportunities for emergence into productive professional careers in between, is useful for purposes of orientation. It allows, for example, the creation of new degree programs specifically designed to prepare future college or high school teachers (see some of the discussion in Chapter 2).

New graduate programs should be established only if they are based on fresh educational concepts and not if they merely mimic and dilute the existing pattern. It may be hoped that more and more well-trained physicists will find teaching positions in colleges and secondary schools. These faculty members could remain in touch with educational and subject matter research through the establishment of regional networks with focal points in strong neighboring university departments. Such regional consortiums would allow teachers to participate in summer programs, weekend conferences, faculty and equipment exchanges, and released time programs, and they would afford established graduate departments a new sense of contributing to the common good.

This is a time for self-renewal. The burden of change lies with the academic community as it persuades the public that the physicists with advanced training are willing and able to interact with society and prepared to contribute as professional physicists to the solutions of contemporary problems. The myth of the "overtraining" of physicists with advanced degrees must be effectively combated, and the professional outlook of our graduate students must be altered to encourage a greater willingness to see excitement and adventure in careers that are not directly oriented toward academic work and research. It is more than likely that physics itself will benefit from the increased contacts with cognate areas. If it takes advantage of the opportunities inherent in the present turmoil, graduate education in physics will emerge even stronger than it is now.

9 The Institutions of Physics Education

In discussing the institutions of physics education, there are two possible points from which to start: with the schools and colleges responsible for the initial education of the physicist or with the various organizations that play a role in his continued development and professional life. The latter category is the main topic of this chapter, for it is widely assumed that a school's responsibility for its product ends with graduation, except perhaps to dun him for alumni dues or support of its athletic program.

The tendency toward the abrupt severance of relations between the developing teacher and his undergraduate college or his university is gradually weakening, however. The trend is likely to continue as emphasis shifts toward quality of preparation. If the institutions preparing physics educators maintain a continuing involvement with them, then the priorities of the scientific and education societies will require re-evaluation. The approach of this chapter is to consider first the role of the educational institutions and then that of the professional organizations.

9.1 THE COLLEGE AND THE TEACHER

Until fairly recently, at least, most prospective teachers of scientific disciplines received the "subject matter" content of their education in a traditional lecture mode with no available alternatives. The key to success—a good grade—lay more in mastering the instructor's examining techniques than in understanding the subject matter. At the same time, the typical would-be schoolteacher pursued the pedagogical arts under the tutelage of a completely different faculty, but generally one with an equally formal and rigid style. At no time did the professional partners in this training enterprise cooperate actively for the benefit of the student; indeed, the two faculties often regarded one another with mistrust, if not downright disrespect. The luckless student saw the two parts of his preparation coalesce only when he reached the ten- or fifteen-week period, at the very end of his college career, when he embarked on trial teaching under supervision of an experienced teacher. At least this has been true of the prospective high school teacher, who presently receives an insight denied his colleague planning to teach at college level, where a background in pedagogy is gained only through experience.

It is not intended to lay blame on one group or the other or to attempt an historical summary of the factors that are producing a change today. It is adequate simply to state that attitudes are changing. Science and education faculties in growing numbers are acknowledging a joint responsibility for good teaching. There is one critically important corollary to be stressed, however. If school and college teachers are to be cognizant of and responsive to the differing needs of their various pupils, they must be able to respond with a range of materials and teaching strategies. No longer will the selection of a single textbook and a list of ten experiments suffice. At the very least, the teacher must be aware of a wide variety of curricular materials and pedagogical techniques and be able to adapt them to individual needs. At best, he will satisfy his drive for creative activity by devising some of his own materials, methods, and evaluative procedures.

Practicing teachers must have continuing, convenient access to the latest curricular materials and established pedagogical techniques. Audiovisual centers, curriculum libraries, and computer facilities should be readily available for their use. Since these are resources that must play a vital role in a teacher's original training, it seems appropriate for institutions that prepare teachers to make resource centers available after graduation. Ideally, they should be available to all teachers in the vicinity, not only to the institution's graduates. Nor should the centers be restricted to the passive role of assembling collections; rather, they should be responsible for offering in-service workshops, seminars, and intensive summer programs. It should be a normal part of a teacher's activity to use resource center materials and to participate in its programs. Finally, it must be stressed that these facilities and programs need the involvement of all departments at the host college, not only the education department. To be most effective, centers must be interdisciplinary in scope.

In recent years, some of the nation's leading universities have established educational research centers to assist their own faculties with a wide range of pedagogical research and development activities. This is a commendable trend, and it is desirable that it spread widely among institutions with a serious commitment to understanding the learning process. Since few smaller institutions will be able to support substantial research centers of their own, those at the larger universities should attempt to meet the needs of the region around them. This might be done through regular distribution of research reports, sponsorship of regional conferences, and establishment of cooperative research projects with qualified faculty from other institutions. Expert advice and the latest research information on educational developments should be readily available to all interested teachers.

In addition to the two kinds of institutions already mentioned, the resource center and the research center, a third category is needed, the major curriculum development center. Its unique function would be to prepare, test, revise, and license for market a broad range of widely useful curricular materials. Only a few such centers in the sciences appear necessary throughout the country. To the extent that they serve national needs and thus deserve extensive federal support, they should be looked on as national centers, completely analogous to the national research laboratories. They might be administered as part of a single university, by a consortium of institutions, or as an autonomous organization.

9.2 PROFESSIONAL ORGANIZATIONS

The professional interests of physicists in education have been widely represented by three organizations, the American Association of Physics Teachers (AAPT), the American Institute of Physics (AIP) through its Division of Education and Manpower, and (until 1971) the Commission on College Physics. In addition, a few of the other professional societies, notably the American Astronomical Society, the Optical Society of America, and the Acoustical Society of America, maintain strong educational programs, although smaller in scope.

The AAPT is a voluntary-membership organization of over 13,000. Its members come from college and university faculties of all kinds as well as from high school faculties. Typically, it has confined its programs in the past largely to holding meetings and publishing journals. An array of committees has performed special projects such as the simulation of laboratory improvement, the production of films, and the recognition of outstanding achievement through the bestowal of awards. Currently, circumstances strongly suggest that the AAPT should assume still broader responsibilities.

The AIP is a federation of professional societies of physicists and astronomers, created as an operating agency to perform functions in the common interest of the member societies. Its Education Division has been engaged principally with the collection and dissemination of data and information. It has established the Niels Bohr Library, including a History of Physics Program, which over the years has accumulated a unique collection of materials of great value to teachers and scholars. In addition, with funds obtained through grants, it has operated a visiting scientists program and a consultant program that have supplied short-term visitors to smaller or remote institutions seeking help with their curricula and the stimulation of contact with other physicists. Also with grant funds, the Division has undertaken the technical physics development program mentioned in Chapter 4, which will be continued under AAPT. With funds derived from its own activities, the AIP supports the Society of Physics Students. But financial pressure has caused the management of AIP to curtail expenditures not supported directly by grants from outside agencies, and the staff of the Education Division is being reduced drastically. The Niels Bohr Library and the History of Physics Project are threatened with an almost complete halt.

Concern with physics education is not confined to those who are members of AAPT, nor to those who are teachers at whatever level. Of the 49,000 individuals represented by AIP, less than one third signify their concern through membership in AAPT. For most of the larger group, programs undertaken by AIP are the only organized expression of their stake in improving physics education. The member societies of AIP need to reconsider their interest in preserving, strengthening, and enlarging their commitment to an AIP education program, especially at this time.

In the 1960's, a new way of promoting science education came into being, the College Commission. Each of the basic sciences and mathematics and engineering established semiautonomous bodies charged with the task of stimulating improvement of college-level education in their fields. All of these bodies were supported by the NSF. The 1970's have brought their demise, apparently as a result of a Foundation policy decision not to be the sole support of any long-lived institution and of wariness of measures that increase the attractiveness of scientific careers. The Commission on College Physics, therefore, was officially terminated in 1971. Its decade of existence had seen many achievements, including the establishment of independent regional organizations, the stimulation of new curricula, textbooks, and films, and the convening of conferences and dissemination of reports. Perhaps none of its accomplishments was more valuable than the education of hundreds of physicists in modes of action designed to bring about educational reform. The value of an independent, semiautonomous central group charged with responsibility to discover needs and stimulate others to act on them is one of the greatest lessons of the decade in science education. The means by which such a group can be made sufficiently representative of and responsive to the community, yet sufficiently detached to exercise critical judgment, and empowered to act promptly remain difficult problems, as they have throughout history.

In its final months, the Commission attempted to maintain a forward-looking stance by organizing two series of studies. One of these was a series of regional conferences on graduate preparation for teaching that attempted to stimulate students and their departments to devise and mount programs to give the teaching assistant guidance and practice more befitting his status as an apprentice teacher than was customary. The second activity was a series of preparatory meetings and a final conference, Priorities for Undergraduate Education, held during the summer of 1971. Although operating from a different point of view from that of the Physics Survey Committee, the priorities conference advocated many of the same changes in attitude and practice. The final Commission Newsletter[57] gives a summary of the conference. The complete report and the Final Report of the Commission, to be published in 1972,* will be significant documents for anyone concerned with physics and education.

To continue the Commission's function of stimulating educational improvement in physics, AAPT has enlarged its staff and has created a new body, the Council on Physics Education. The mandate of the Council directs it to maintain surveillance over the full range of educational levels, recommending and promoting action wherever needed. The membership of the Council is more widely representative than was that of the Commission, in recognition of its broader responsibility. More directly responsible to a membership organization than was the Commission, it must find a mode of operation combining this responsibility with prompt and discriminating action. If the Council is to be broadly effective, it should have working relationships with other organizations with interests in physics education. The AIP has been invited to appoint a representative to the Council; other societies might well be included in this formal manner. It was necessary for the AAPT to underwrite the cost of staff increases and Council operation from its own resources, which necessitated a significant increase in dues. The membership of the AAPT has responded favorably to the added burden.

9.3 THE LOCAL SCENE

So far, this discourse has focused on two elements of the organizational pattern in physics education: the educational institutions and the professional societies. Stated briefly, the institutions provide the teacher with his formal studies; his methods and materials, both initially and as a practitioner; and, at least ideally, with a resource center and an active research and development laboratory. The societies offer the teacher a variety of information in several modes, afford him entrance to those same channels of communication, and, hopefully, take stands on issues of importance and interest to him.

In one sense, both the educational institutions and the professional societies offer the same commodity–information. The institution centers will emphasize materials and techniques particularly appropriate to the instructor's level of teaching. Their services will be offered equally to all disciplines. The professional societies, on the other hand, sponsor activities that cut vertically across all levels of teaching within a single discipline.

* According to Newsletter No. 25, these reports were to be published in the *American Journal of Physics* during 1972.

The two classes of organization interact in the sense that the society can identify problems of national importance and recommend to various institutions that solutions be sought; the results can then be communicated to the individual teachers for their consideration and possible implementation. It is at the local level that the efforts of both organizational types come to fruition, because it is in the individual classroom that changes are made. Thus, brief consideration will be given to the activities of local and regional organizations before more specific recommendations are formulated for the various national societies and federal agencies.

9.4 REGIONAL ORGANIZATIONS AND CENTERS

One of the more far-reaching accomplishments of the Commission on College Physics was the conception and nurturing of a number of regional organizations intended to conduct localized efforts to improve physics teaching. The six that have evolved[58-60] are diverse in constitution and programs, reflecting the quite different circumstances that brought them into being. In some cases, they bring together a fairly homogeneous group of colleges that can easily share instructors and facilities; in others, they encompass the full range of educational institutions, from two-year technical institutes to universities, public and private. Organizations of this kind can apply for grants to support their activities; in at least one case, the approximately 60 member institutions pay dues that partially support a central office and newsletter. Among the explicit aims of such groups, the following are prominent:

To increase communication and cooperation among regional physicists. (In some parts of the United States, a college physics instructor has been known to spend as much as five years on the faculty without any personal interaction with his colleagues.)

To offer instruction to physics teachers with inadequate or outdated academic training.

To offer local opportunities to hear about new instructional materials and developments from those engaged in producing them.

To offer opportunities for instructors to discuss their problems and ideas with others in similar situations, often guided by more experienced and highly trained physicists.

Experience has shown that, even when funds have been difficult to obtain and programs have had to be curtailed, regional organizations maintain a vitality that could result only from their meeting an important need—

the need for a stronger feeling of community than that derived from non-discipline-oriented associations or from the more diffuse and conventional professional organizations. Although aspects of their programs may have a familiar appearance, their role and effectiveness must be judged in a context that differs from the more usual centralized frame of reference of funding agencies.

There can be no doubt that informal communication with his colleagues, whether they come from within the same metropolitan area or are drawn from throughout a state or perhaps even from a multistate region, is the single most valuable experience afforded a teacher by his local organization. Interest in such a group can be effectively stimulated by a regular, concise newsletter; it can be sustained by occasional meetings at which formal or semiformal presentations are made; but it is really most fruitful in the private conversations, campus visits, and working relationships that follow. No communication is quite so effective as the conversation that follows the question, "How do you handle this problem at your institution?" or "What teaching strategies do you use to present this topic to your students?"

Another very important function of a local society, one that regrettably is served much less frequently than the local communication role, is to bring nationally recognized experts to the local scene to share their knowledge of particular developments in physics education. Whether via lectures, discussions, workshops—whatever the mode of presentation—this constitutes local implementation in a highly stimulating form. Furthermore, the individuals who introduce new techniques and materials as a direct result of the visitor's presentation are aided in their efforts by the continuing communication channels of the local group, and they exert a very strong influence on their less creative colleagues to try something new, too.

A few local organizations have played a third very helpful role by arranging cooperative use of specialized research facilities at leading institutions by faculty and students from other institutions in the area. Some have promoted exchange of equipment that has been discarded by one department but is still valuable to another. Sharing of resource collections and cooperative use of educational research centers seem particularly valuable objectives for local groups at this time. Indeed, they might be effective in promoting establishment of such centers in regions where none now exist. Faculty exchange between institutions of differing types has often been suggested, but there is little evidence of successful consummation of such an arrangement.

Finally, local groups, like national organizations, can take stands on important issues and attempt to change attitudes. While it is true that their effectiveness may be enhanced by gaining national endorsement, it is equally true that they can serve an important function by pointing out problems of local concern that might benefit from national attention.

9.5 OTHER INSTITUTIONS

Several other AIP member societies, as we have mentioned, also have active education programs; these include the Optical Society of America, the Acoustical Society of America, and the American Astronomical Society. Beyond these, the American Society for Engineering Education has a division on physics education that shares both members and interests with the physics societies. Such groups as the American Association for the Advancement of Science, the Cooperative Committee on the Teaching of Science and Mathematics, and the National Science Teachers Association conduct programs that include physics as one of many disciplines. There are many additional organizations with education programs in mathematics or one of the sciences other than physics. It might also be worth mentioning such groups as the American Council on Education and the Associated Organization for Teacher Education, with both of which there is clearly mutual interest on a more general level.

It would obviously take someone with the wisdom of Solomon and audacity to match to carve out unique areas of operation for each of the organizations we have mentioned. Lacking both omniscience and omnipotence, we can only urge that they stay in close touch. Where feasible, meetings of the principal educational committees should be attended by observers from other interested societies; where possible, societies should seek formal affiliation with coordinating groups. Communication and cooperation should always be stressed. More specifically, the Council on Physics Education should be encouraged to expand its practice of inviting observers from other organizations in physics or closely related to physics to sit in on its deliberations. This is but a first step toward the very desirable goal of establishing such exchange arrangements on a permanent basis among leading science education councils. The practice of appointing to CPE membership individuals who can represent not only physics but also a society other than AAPT is also commendable and to be encouraged. The development of a high-level, multidisciplinary science education council, structured like the Cooperative Committee on the Teaching of Science and Mathematics, but with a membership much closer to the seats of individual society power, is a worthy long-range objective.

Federal support of science education is a vitally important concept and is treated in the following chapter. The few general remarks added here are largely in the context of the previous discussion of institutions and programs. Considerable rejuvenation of high school and college science curricula has been achieved through major national curriculum development projects. These should be continued, particularly as they pertain to nonscience majors and to creation of more effective and realistic methods of instruction. They must not be aimed only, or even primarily, at the bright students who will succeed no matter what the challenge. To be successful, such projects must include a mix of competent scientists and experienced educators in their directorates. The National Science Foundation has been particularly effective in encouraging these characteristics among its grant recipients.

Another area to which increasing attention must be given is that of practical educational research. By this is meant study of the learning processes and investigations into the merits of various teaching strategies. Attempts continuing and extending those mentioned in Chapter 1 should be made in order to determine the effectiveness, in terms of demonstrated student learning, of major curriculum developments. Major results of such studies should be disseminated widely to stimulate more productive teaching. We believe it would be well for scientists to work with their education colleagues in these studies.

One area in which funding activity has been particularly noticeable by its almost complete absence is the stimulation of effective local implementation. In general, the results of curriculum development projects have been put in small packages on bookstore shelves to be marketed like any other commercial product. To be sure, summer institutes and in-service programs have been built around these packages, but not always with leaders experienced in their use and seldom with any continuing long-range interaction among participants. Rarely is a new program developed with the idea that instructors will select only small parts of it, to be modified and joined with elements of other programs into a unique course suitable for the unique needs of particular students. Rather, an instructor is expected to choose all of one, none of the rest, and to follow it faithfully. No wonder so many individuals teach by reciting glibly exactly what is in the book.

What is so desperately, and obviously, needed is a series of "permanent" programs of rejuvenation and creativity, in which all practicing science teachers participate regularly as a part of their normal work. The view respecting this need that is held in the Soviet Union has been quoted in Chapter 2. These programs could appropriately utilize institutional resource centers of the kind mentioned previously in this chapter. Some portions of these activities will be organized by the host faculty, with its expertise and experience in many disciplines, but others can more effectively be arranged by interinstitutional groups embracing a single discipline. The effectiveness of these latter groups in generating productive change has been almost completely ignored, probably because so little money has been granted to them. In the area of physics, consortia like the Chesapeake Physics Association, the Illinois State Physics Project, the Austin Peay

State University Center for Teachers, the Pacific Northwest Association for College Physics, and the several "COSIPC groups" come to mind. These all involve physics teachers from various types of institutions within a single, logically defined region in a continuing series of programs designed to meet the expressed needs of the teachers.

10 The Support of Physics Education

For about a decade now the key words with respect to the system of formal education have been criticism, crisis, and change. On a worldwide scale and at all levels, our educational systems have been increasingly called into question. Curricula have come under fire for being too rigid and too narrow, for lacking relevance, for suppressing creativity and imagination, and for isolating students too long from concrete experience with practice and with society. School systems have been perceived as mechanisms for the preservation of the status quo and as a way of denying members of minority groups practical access to desirable levels of society. The economic reverses of the past several years and the increasing difficulties that highly trained men and women have had in obtaining suitable employment have brought charges of lack of realistic vision and planning in higher education. The rapidly rising costs of formal education have put many institutions, both public and private, in extremely precarious financial positions and have caused many legislative bodies to concern themselves with questions of fiscal accountability, efficiency, and productivity in schools and colleges. Even the long-held idea of subsidized public higher education as a positive social good—of as much or more benefit to society as to the student—has been questioned and has begun to wane under increasing economic pressure.[54,61-77]

Physics education, whatever its special aspects, is deeply embedded in the general system of education, and at each level it shares the system's problems. It is in this general context of educational and social flux that the problems of the support of physics education present themselves to us. It seems useful, therefore, to remind ourselves briefly of what the educational system looks like: its size and complexity, and its connections with the other components of the even larger and more complex system of society of which it is an integral part.

Figure XIII.4 is a schematic diagram of the formal educational system and some of its connections to itself and to other systems in society. The

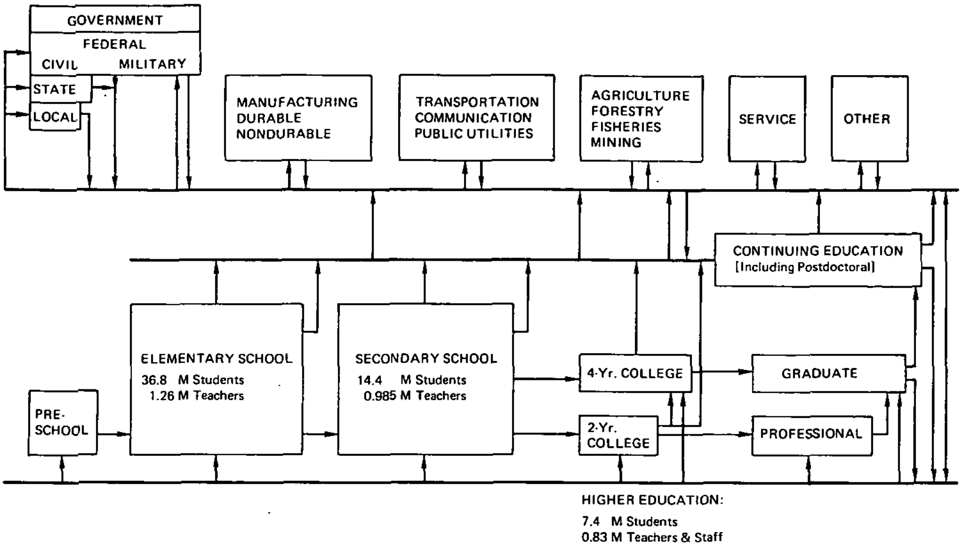


FIGURE XIII.4 The formal educational system (1969–1970) and connections with other systems in society.

statistical data in the diagram provide a measure of the system's size in 1969–1970; the 62 million full-time students, teachers, and administrators constituted about 30 percent of the U.S. population, and expenditures of \$70 billion for the system represented about 7 percent of the gross national product for this period. Figure XIII.5 illustrates a portion of the scheme that relates directly to physics. Figure XIII.6 serves to remind us of the dramatic growth of education during recent years.

10.1 TODAY'S PROBLEMS

Education, people say, is an example of a "service industry." Communications, accounting, and medical practice also are classed as service industries, but they seem to be examples of a different category, to which quantitative measures of "productivity" can be applied and a "unit output" derived. We will not argue that these pursuits lack qualitative values, but in the case of education, the measure of productivity would seem to be primarily qualitative, and a unit output is difficult to define. This difficulty means that support for education cannot be linked to a measure of productivity. During recent years, while the cost of education has risen along with all

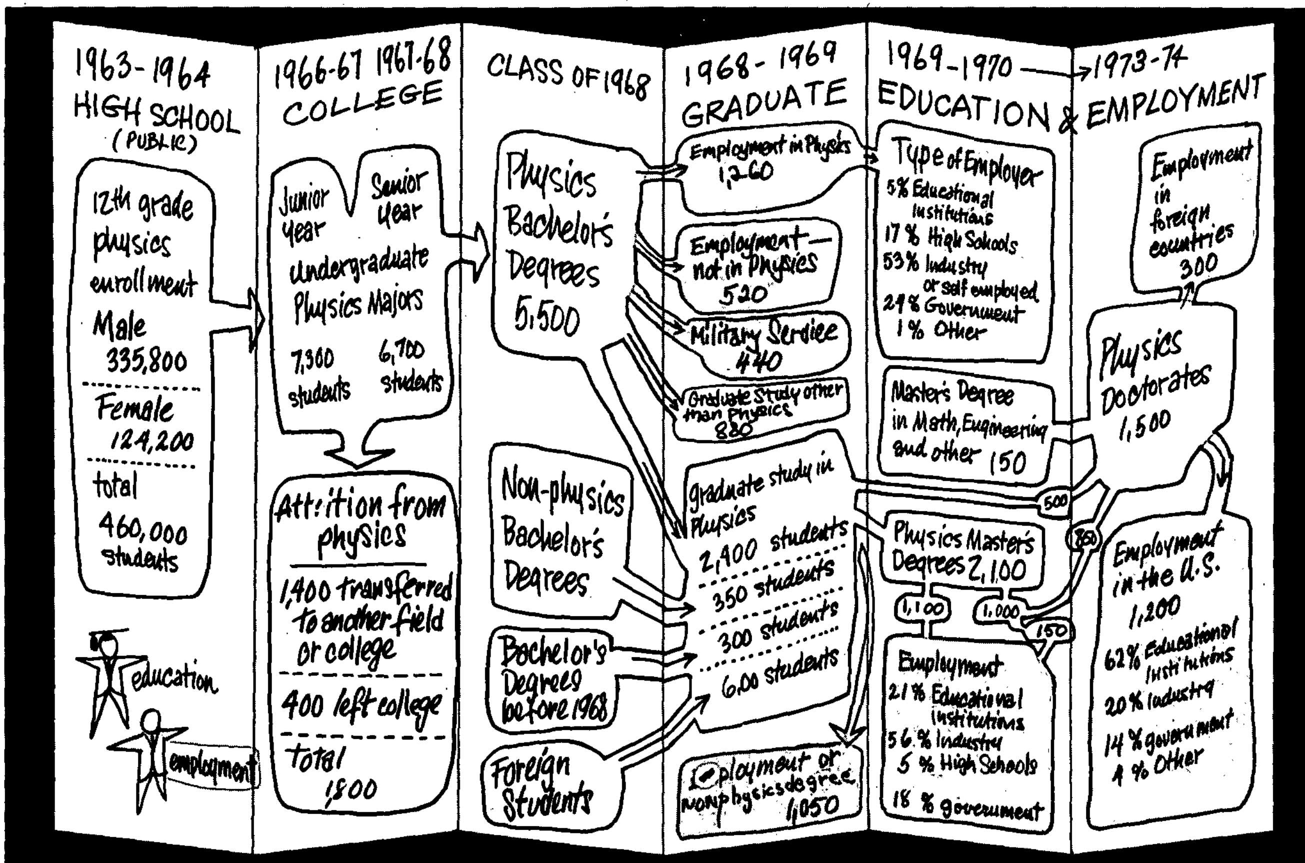


FIGURE XIII.5 Physics education and employment, 1964–1974. [Source: *Physics Manpower 1969*. American Institute of Physics Publication No. R220.]

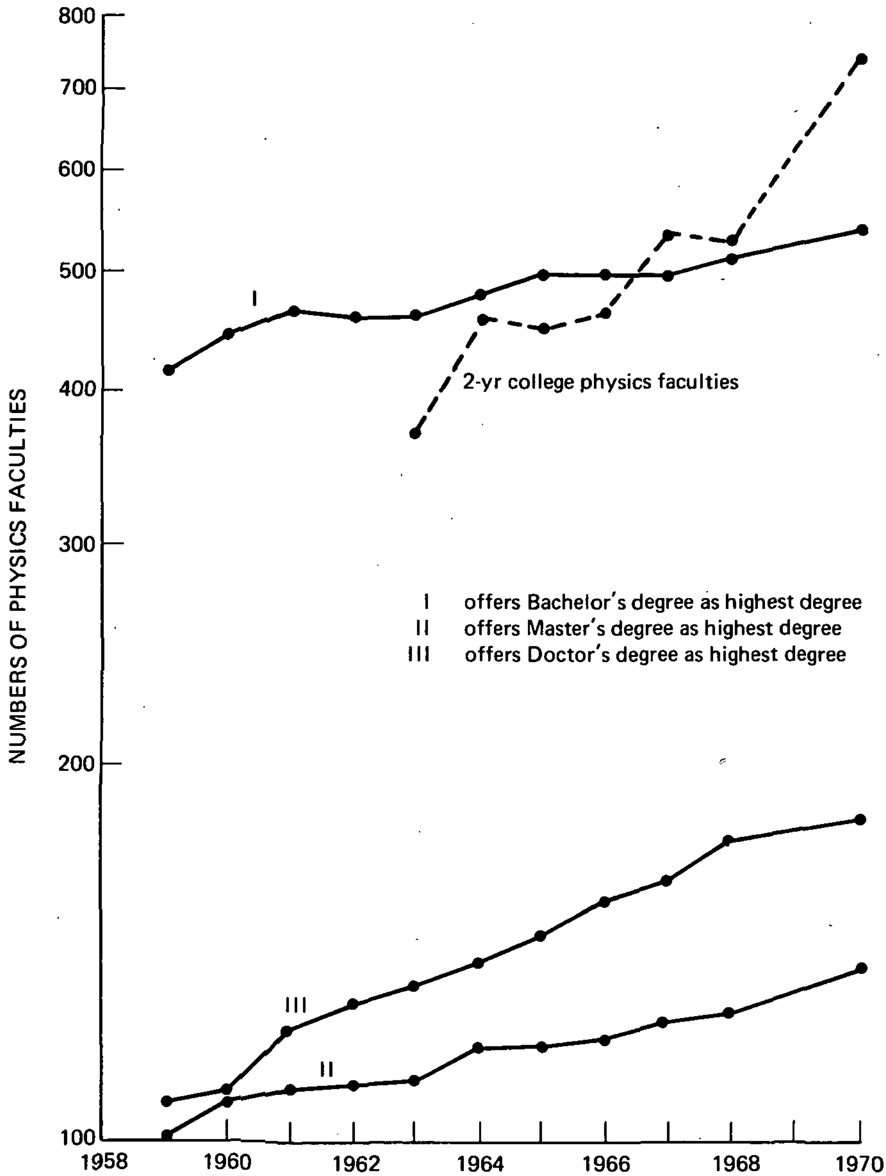


FIGURE XIII.6 Numbers of physics faculties listed in AIP directories, 1959–1970. [Source: AIP physics manpower and education statistics.]

other costs, there has been no way to demonstrate an increase in educational productivity. The cost of education grows more prominent in the national economy, and support for it finds less favor. For this reason, if not for others, educational institutions at all levels are facing critical economic problems at the present time, and some are succumbing to the pressure. This crisis lies at the base of all discussions of educational support.[78]

For purposes of making economic and manpower forecasts, attempts have been made to construct quantitative models of educational systems (see, for example, Blang[79]) and even to incorporate into them factors as important but as difficult to quantify and to measure as social values (e.g., Baier and Rescher[80]). To the extent that such models have been realistic, they have also been complex and have led for the most part to an increased understanding of the difficulties associated with model construction and with obtaining relevant and accurate input data.[79,81] Nevertheless, the model approach can be useful in efforts to visualize the structure of complex systems and to comprehend the nature of their dynamics. In a partial and qualitative sense at least, Figure XIII.5 may also serve in this respect.

A prominent feature of Figure XIII.5 is the number of loops or internal cycles it contains, a manifestation of its complexity. We have learned to ask questions about the properties that such complexity confers on a system, and one of the most significant questions is about the capability of the system for self-renewal. In his book *Self Renewal,* John Gardner[82] makes this point more directly: ". . . Over the centuries the classic question of social reform has been, 'How can we cure this or that specific ill?' Now we must ask another kind of question: 'How can we design a system that will continuously reform (i.e., renew) itself, beginning with presently specifiable ills and moving on to ills that we can not now foresee?' " "In the ever-renewing society," says Gardner, "what matures is a system or framework within which continuous innovation, renewal and rebirth can occur."

The many efforts to refresh physics education that we have described in our report, as well as those we have not mentioned explicitly, are indeed responses to perceived ills, and, taken as a whole, they amount to a considerable potential for self-renewal. However, for the most part they are the consequence of individual enterprise, the response of sensitive persons to social needs they have perceived. From the larger perspective of this chapter, self-renewal needs to be included as an important, visible, institutionalized feature of the system. The Panel on Educational Research and Development, in its 1964 Progress Report,[83] put it thus:

The effort to improve education—to develop better curricular materials, better programs of teacher education, better schools and school systems—is not a one shot affair. This activity should be carried on continuously. At the heart of the current effort lies the assumption that nobody knows the "ideal" system. Meeting immediate needs can prepare the way for longer-range reform, and the new results in fundamental research will open up new possibilities. *Changes in the schools will make possible changes in the colleges, and changes in the colleges will make possible changes in the schools.* [Emphasis supplied.]

The curriculum projects mentioned above and programs such as NSF's Pre-Service Teacher Education Program (UPSTEP) and Cooperative College–School Science Program (CCSS) are examples of useful components of a self-renewing system, but as yet they affect only relatively small percentages of the 34 million pupils and 1.3 million teachers in the elementary grades. Our educational system would seem to be very far from maturity in Gardner's sense.

One other aspect of support requires additional comment at this point—namely, the role of the federal government and federal agencies such as the Office of Education and the National Science Foundation. About 93 percent of the $46 billion cost of elementary and secondary education in 1970 came from nonfederal funds. But federal support can exert and has exerted important catalytic effects. For example, the investment in primary and secondary school course content improvement projects by the National Science Foundation averaged over the period from 1957 to 1969 comes to only about four one-hundredths of 1 percent of the corresponding national operating cost for education at these levels.

At least until the educational system has matured to the point where it is generally recognized that the small percentage costs of continuing educational research and development should be included in operating budgets at all levels of the system, the availability of continuing federal funding in this area is likely to be crucial to the progress of educational reform. From this point of view, it is discouraging indeed to note the progressive decline in federal appropriations for such education programs.[84] Figure XIII.7 illustrates a trend in federal support that has been criticized by the House Committee on Science and Astronautics.[85] There are indications that the trend may be reversed, in the discussions of the NSF budget now taking place and in the legislation introduced for the establishment of a National Institute for Education.

10.2 ACCOUNTABILITY, COST EFFECTIVENESS, PRODUCTIVITY, AND EDUCATIONAL TECHNOLOGY

In the two decades from 1950 to 1970, the annual cost per pupil, adjusted to a constant dollar basis, increased from $377 to $885 for elementary and secondary schools. For higher education, the corresponding increase was from $1219 to $3112. For the past several years, large numbers of private

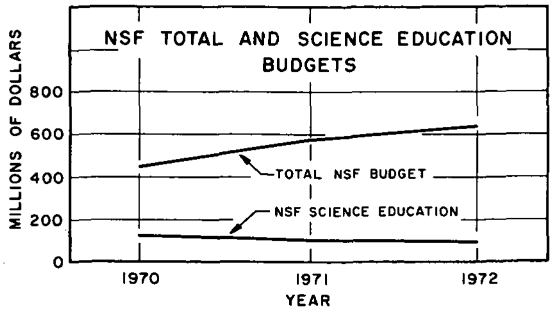


FIGURE XIII.7 Total yearly NSF budget and yearly NSF budget for education. [Source: Commission on College Physics.]

schools, colleges, and universities have been running on deficit budgets and are now near the brink of financial disaster. A substantial fraction of local school levies fail at the polls with dismal regularity, and recent appropriations for higher education by state legislatures have been largely at standstill levels or less.

Students, their parents, school boards, boards of trustees, state legislatures, federal agencies, foundations, and the business community, all have been questioning the programs and performance of our schools and colleges; and such questions are being asked not only in this country, but worldwide.[1] A recent article[86] on the Open University in Britain concludes with the remark: ". . . what the Open University has done is to develop a teaching *system* that is potentially applicable to any country; and on the scale on which it has operated in Britain, it will produce a graduate at about 20 percent of the cost of a conventional university graduate. In these cost-conscious days this is its source of interest to the world's governments."

But it is not only the total cost of education that is at issue. The Newman Report[87] for example, puts it so:

> . . . Thinking about costs is not simply a matter of paring budgets and making ends meet, of cutting out secretaries or not buying typewriters. It is a fundamental educational issue. Searching for more effective methods of teaching must lead us to examine the neglected questions of what we are trying to do and how students learn.
>
> For the university as well as for society, the issue is effective use of resources. If time and energy can be saved by adopting more cost-effective procedures, those energies can be devoted to a long list of tasks now starved for resources. . . . Considering what needs to be done, we can afford the high cost of education, but not the low productivity.

Questions of efficiency and productivity in the educational system have been with us at least since the turn of the century.[88] Their present return to prominence has many causes, one of which is related to the increasing shift in employment from the production of goods to the production of services. In an editorial in *Science*,[89] Patrick Haggerty suggests that

> . . . the plentitude of industrial and consumer goods we now produce, plus all mining and construction, require about the same one-third of our work force as did the relatively much lower output of 1890. It therefore has been possible to increase the number of those working in the service sector from about 25 percent in 1890 to above 60 percent today. That shift has been made both because we could make it and because we wanted to do so. We wanted more health services and more people in education. We wanted to move to the cities, with all that implies in the way of government and governmental services. . . . This shift in effort from goods production to services production results in a constantly increasing proportion of our total work force being engaged where, in general, we simply have not learned how to be as effective as we are in goods production.

In discussing the question of productivity in industry and in education, it may be useful to remind ourselves that we are now using the term in a more conventional sense than we did in Chapter 3.

The difficulties in defining productivity, or effectiveness, or inputs and outputs, in an educational system—let alone the problems of their quantitative measurement—are obvious and formidable, but they need not and should not deter us from attempting to better understand the nature of the educational process and from attempting to make it as "effective" as we can. If one is looking for order-of-magnitude effects, or even factors of 2 or 3, then precision in analysis and in definition may not be of immediate importance.

Many proposals have been made of ways of increasing productivity in education, some dealing with support functions and others with the educational process itself.[87-90] There are, for example, the possibilities for economies of scale. Presumably, increasing the size of an institution can lead to reduced unit costs and to the attainment of the "critical mass" of scholars in a given field requisite for a high degree of academic excellence. One of the most recent studies of size effects is that of Gallant and Prothero.[91] Their analysis of data from the California system leads them to the conclusion that unit costs of education in a university decline very little above a size of 10,000 to 15,000 students, and that continued growth can in fact lead to dysfunctional consequences that tend to reduce the overall effectiveness of the institution. "It is difficult to see," say these authors, "what further advantage, other than the possibility of United Nations membership, can accrue to a university population above 10,000 souls."

Interinstitutional cooperation is another possible mechanism for reducing costs, and many formal structures for facilitating such cooperation exist—for example, the Committee on Institutional Cooperation, whose members are eleven major midwestern universities. In practice, however, the cooperative activities among these institutions appear to be quite small compared to their individual activities, a situation in which, no doubt, the factor of geography plays an important role.

Over a century ago, T. H. Huxley, in a talk on scientific education in the schools, noted that "Under the regulations to which I refer, a schoolmaster can set up a class in one or more branches of science; his pupils will be examined, and the State will pay him, at a certain rate, for all who succeed in passing." With the idea that the private sector, with its incentives to efficiency, might produce results substantially better than those of the public schools in the teaching of reading and elementary mathematics, performance contracting has come into public prominence within the last couple of years—some contracts even with the "no pass/no pay" provisions of Huxley's time. At the end of the 1970–1971 school year, some 54 school districts in 24 states had participated in performance contracting experiments. Individualization of instruction, teaching machines, special materials, and material incentives for students and teachers were used in various combinations. The analysis of these experiments is at this time still in a preliminary stage, but ". . . these results . . . suggest that some of the programs did worse than conventional methods, most did about as well, and a few did a little better. The initial indications were that experiments stressing incentives did better than those stressing teaching machines."[92] This apparent inability of the private sector to substantially better the results of the public schools on a large scale no doubt has multiple explanations, but an important one among them most likely is overestimation of our present understanding of how the process of learning works in individuals, and of how, successfully, to mediate it.

In the production of material goods, major advances in productivity have resulted from such factors as quality controls on the raw materials used, the introduction of techniques and management systems for mass production, the introduction of technology in the form of machines which compete effectively with human labor, and the invention of radically new processes of production.

Much has been said and written about the potential of technology for extending educational opportunities very widely and at low cost.[93] Indeed, there is historical precedent for this. The introduction at about the middle of the fifteenth century of what is still the single most common technological tool of education, the printed book, clearly radicalized the whole process of education and also made possible self-instruction in a way and to an extent that had not previously existed.

Modern communications technology, in the form of telecasts, cable television, video cassettes, cartridged film loops, and other audiovisual media, is seen by many as having at least a similar potential. For example, "Sesame Street," the educational television show for children, has provided a 26-week, 130-program series to an audience estimated at seven million. The cost of producing this series in fiscal 1970 was about $6.5 million. If the estimate of the size of the audience is accurate, the unit cost of production was less than one cent per child per hour.[93] Of course, the "student terminals" for such television broadcasts already exist in over 98 percent of the households in this country, something over 40 percent of them being color TV receivers.

Entirely aside from questions such as how one evaluates the "effectiveness" of programs of this sort, it is of interest to note that their relatively small per-unit costs can be achieved only by having a huge mass audience for the product. There is no question of accommodating such instruction to individual needs; all seven million children see and hear the same videotape product, and, as with a book, the communication is necessarily one-way.

The use of computers in educational support functions (in administrative and other data-processing activities) and in scientific computation is too well known to require comment here.[94–97] Their use as major components of instructional systems has also been explored and will be taken up in Chapter 9. Developments in this area have been slow and expensive although there have been substantial decreases in the costs of certain kinds of computer hardware since 1969.

Costs per student-hour of instruction utilizing modern technology have been examined by Esseff,[98] some of whose results are reproduced in Table XIII.4. He notes that "Computer-assisted instruction has been omitted since the costs are extremely variable depending on the nature of the installation and its operation, although these total costs typically run from $2 to $80 per student-hour." There have been predictions for a decrease in costs for certain special kinds of computer-assisted instruction to about $0.50 per student-hour, but this has yet to be achieved in practice.

In his monograph for the Carnegie Commission on Higher Education, William G. Bowen comments as follows[99]:

> . . . Considerable attention is being given to the development of teaching aids of various kinds, and everyone concerned about education can only hope that ways will be found to teach more effectively at lower cost. However, in speculating about the significance of these developments for cost-per-student in the decade ahead, it is well to recognize that many of the most exciting ideas are still untested, and that even if they turn out as well as is hoped, practical implementation on any large scale would require considerable lead time and administrative effort—as well as substantial capital expenditures.
>
> An even more important question from the perspective of the institutions with which we are primarily concerned in this paper is whether electronic teaching aids are

TABLE XIII.4 Costs per Student-Hour of Major Instructional Systems[a]

System[b]	Total System Costs ($)	Cost Sensitivities[c]: Decrease on Lines of Maximum Potential Decrease ($)	Cost Sensitivities[c]: Increase on Lines of Maximum Potential Increase ($)	Cost Sensitivity Index	Role of Initial Costs (%)	Role of Operational Costs (%)
CLT	0.28	0.29	0.84	3.45	18	82
LL (List)	0.80	0.60	5.22	1.15	15	85
DAIRS (A)	0.81	0.62	5.36	1.16	23	77
PI (L)	0.84	0.70	6.12	1.14	19	81
PI (B)	0.97	0.78	6.94	1.12	22	78
DAIRS (A/V)	1.33	0.91	8.20	1.11	35	65
VTR	1.40	0.88	0.88	0.11	51	49
TM (L)	1.75	0.57	6.37	0.016	64	36
LL (Lang)	1.99	0.82	6.75	0.015	64	36
CCTV (w/f)	2.37	1.98	18.03	1.10	21	79
SRS (T/S)	2.37	1.81	17.28	1.05	25	75
CCTV (wo/f)	2.43	2.11	19.35	1.09	27	73
PIA	5.50	2.47	35.08	0.70	33	67
LCD (Snd FS)	6.56	−1.45	29.84	0.003	61	39
TM (B)	7.42	−6.90	16.03	0.006	88	12
SRS (A/V)	8.90	5.17	68.41	0.76	24	76
LCD (Snd FS cart)	16.25	−9.82	47.04	0.002	84	16
LCD (Sil MP)	43.40	−19.04	170.88	0.0006	68	32
LCD (Snd MP)	59.92	−39.23	195.81	0.0002	76	24

[a] Data from Esseff.[98]

[b] CLT: Conventional lecture with textbook
CCTV (wo/f): Closed-circuit television without feedback
CCTV (w/f): Closed-circuit television with feedback
PI (L): Linear programmed instruction text
PI (B): Programmed instruction text with branching
TM (L): Linear program teaching machine
TM (B): Teaching machine with branching
TM (B Snd MP): Teaching machine with branching and sound motion picture capability
LL (LIST): Learning laboratory (listening)
LL (LONG): Learning laboratory (language)
SRS (T/S): Teacher–student feedback, student response system
SRS (A/V): Audiovisual control, student response system
DAIRS (A): Dial access information retrieval system (scheduled audio)
DAIRS (A/V): Dial access information retrieval system (scheduled audio/video)
DAIRS (R): Random dial access information retrieval system
PIA: Portable instructor aids
CAI: Computer-assisted instruction
LCD (Snd MP): Learner-centered audiovisual device—sound motion picture
LCD (Sil MP): Learner-centered audiovisual device—silent motion picture
LCD (Snd FS): Learner-centered audiovisual device—sound filmstrip
LCD (Snd FScart): Learner-centered audiovisual device—sound filmstrip cartridge
VTR: Portable videotape recording

[c] $$\text{CSI} = 10 \times \frac{\text{(Total decrease on lines of maximum potential decrease)}}{\text{(Total increase on lines of maximum potential increase)}}$$

CSI is a measure of the sensitivity of student-hour costs to changes in numbers of students or numbers of hours, a large CSI indicating a favorable condition in this respect.

particularly well-suited to the level of education in which they specialize. It is going to be easier to develop pedagogically and economically viable ways of presenting basic subject matter to large groups of students than to find methods of teaching advanced materials which will develop simultaneously the capacities of highly selected students to make new contributions on their own. The conscientious supervision of a student's independent work is the essence of high-level graduate education, and it is an important element in the undergraduate preparation of highly qualified students. Yet it is hard to see how any significant savings in faculty time are to be achieved here.

Bowen's last point becomes particularly important when one recognizes that the per-student costs increase dramatically as one moves from lower-division undergraduate work to graduate work. There appears to be relatively little hard data on the subject, but ballpark estimates seem to indicate that, averaged over the sciences, the relative per-student costs for lower-division undergraduate, upper-division undergraduate, master's, and doctoral level work are about in the ratio of 1:2:4:10.

In physics especially, the costs at the graduate level are high because of the nature of the experimental equipment required and its tendency to become increasingly complex and expensive. A recent British study[100] suggests that this "sophistication factor" is in fact highest for physics and that per capita equipment costs have been increasing at a rate well in excess of 12 percent per year.

For an institution that has a major fraction of its enrollment at the graduate and advanced undergraduate levels, the implications of these arguments are likely to be rather serious. With its operations heavily weighted in areas deemed intrinsically "labor-intensive," it has little possibility of being able to increase appreciably its *overall* productivity. The present plight of many major universities graphically illustrates this dilemma.

To date, educational technology has not been notably successful in improving productivity in education. But it has done one very valuable thing. To quote Oettinger[96]:

. . . every attempt to introduce technological change into education has revealed how profoundly *ignorant* we still are. We know precious little about the psychology of learning, and what we know is more relevant to the laboratory than to the classroom. Behavioral objectives with a strong taste of the explicit, the quantitative, and the measurable account for but a small fraction of the many effects we expect of schooling. We have dream devices that we cannot make work effectively. We think of reorganizing institutions, but contemporary social and political science, though abounding with static descriptions, tell us next to nothing about the dynamics of transition from one form of organization to another. There is a history of instruction in industrial, military, and more recently, Job Corps settings, but its relevance and applicability to the *formal* education of children is unknown. The enormous educational influence of mass media, peers, and other information agents is obvious, but no one knows how to control and to maximize this influence or how to balance the formal and the informal.

. . . Finally, the notion that any form of technology can make a significant contribution "at no additional cost or even at lessened costs per pupil" . . . is an illusion. More books and better libraries cost more money. Greater individualization costs more money, no matter what the specific process may be. Better understanding and better-trained people cost more money.

10.3 PHYSICS EDUCATION

In its recent report to the National Science Board,[101] the NSF Advisory Committee for Science Education stresses an "overriding goal for the coming decade: To educate scientists who will be at home in society, and to educate a society that will be at home with science." Similar concerns have been expressed by other groups concerned more specifically with physics education,* and this theme certainly undergirds our report. If the goal is to be reached, unprecedented support will be required—in ideas and in manpower, even prior to money.

Referring to Figure XIII.5 and recalling some data from earlier chapters help to illustrate the dimensions of the problems. In round numbers, there are about 37 million students and 1.3 million teachers in 85,000 elementary schools; and about 15 million students and nearly a million teachers in 31,000 secondary schools. There are 2500 colleges and universities, with 7 million students and 0.8 million instructors.[10]

Reliable information on a national scale concerning the extent to which science is taught throughout the system and the scientific literacy of the public is, as we have seen, quite difficult to obtain. Because of the cost and the effort required to produce them, extensive national surveys are made infrequently. Very often, too, it is difficult to compare the results of different surveys because they have used significantly different designs and data collection procedures or different data bases. The point is made explicitly for the secondary schools by Chin[102] that we do not even have current and accurate information on the total enrollment in science and in the various science courses. Maben[103] also comments on this same difficulty with respect to information about science in the elementary schools,

* "Priorities for Undergraduate Physics Education: A Report of the Commission on College Physics Conference" (to have been published in the *American Journal of Physics*, in 1972). A preliminary account has appeared in the Commission on College Physics Newsletter No. 25, November 1971. This issue also contains an account of a Symposium on the Education of Physicists held at the Battelle Seattle Research Center in 1971. The Council on Physics in Education of the AAPT has also expressed similar concerns.

noting that the problems in this area are even more complex than those for the secondary schools. Certainly, one prerequisite for obtaining adequate support is a clearer *understanding* of the existing situation than we now possess, not merely a cataloguing of teachers' credit hours and attitudes, of enrollments, or of classroom equipment. We need surveys that probe more deeply, constructed so that successive ones are comparable. But our report, dwelling on many of the requirements to reach the National Science Foundation's goal, has not rested on a statistical base, nor has it attempted to put forth detailed budgetary or manpower recommendations. We believe that budget and manpower are important matters and that the arms of government—local, state, and federal—responsible for providing support for education would use thoughtful opinions and recommendations of physicists regarding them. However, we believe the most significant advances toward the goal expressed by the National Science Board will come as a consequence of changing attitudes on the part of our college and university faculty colleagues, a large fraction of whom have not yet determined to make the goal their own. The surest route to adequate support is via, first of all, their commitment and action. In all ages, the hallmark of the successful artist, writer, scientist, or teacher has been the ability to invent new and more effective ways of utilizing available resources.

11 Trends and Opportunities

Education, following the strong trends of social evolution, continually opens new vistas. One that receives public attention from time to time is the potential use of technology to increase the opportunity for individualized instruction. The major effort to develop and use new educational methods with special appeal and effectiveness for the culturally deprived has elicited some thoughts on the place of science in this movement. In an era that is questioning formality in any guise, modes of offering college-level education without requiring attendance at a traditionally structured academy are emerging. Many other new trends and opportunities merit attention; the studies of the Carnegie Commission are an important source of information about them. Space permits us to discuss only a few, beginning with the first attempt to learn systematically on a nationwide basis just what U.S. citizens of various age levels know.

11.1 THE NATIONAL ASSESSMENT

Critics of U.S. education often cite the achievement of our youth in tests as compared with that of the youth of other nations. Such comparisons, unless the samples of students are carefully matched in respects other than age or school level, can be misleading, because few nations attempt to provide education to as broad a segment of the population as does the United States. Nonetheless, those in this country who are concerned with educational quality seek some objective standard against which to measure U.S. output.

An absolute standard is not yet in sight, but the National Assessment of Educational Progress* offers an objective measure of change. The project was established as an on-going activity under the Education Commission of the States. It proposes to measure what 100,000 U.S. children (ages 13 and 17) and young adults (26 to 35) actually know in ten areas, including science. Changes in (a) knowledge, (b) skills, (c) understanding, and (d) attitudes are to be assessed by periodic testing of each age group according to a scheme in which some old and some new questions are to be used. The first test including science was held in 1969–1970; the next is scheduled for 1972–1973. A third, in 1975–1976, and a fourth, in 1978–1979, are also scheduled. The results are reported by geographic region, size of community, and sex. They also are reported by area of science and objective (a) through (d) above. The Education Commission of the States proposes to issue reports from time to time that, in addition to statistical data, quote examples of performance but neither generalize nor interpret.

The results of the first assessment have been issued as a report accompanied by a booklet of observations and commentary by a panel of independent reviewers.[104] A statement of particular interest in the context of this report is that nine-year-olds were equally knowledgeable in biology, physical sciences, and earth sciences but that seventeen-year-olds gave a disappointing performance in physical science. Another reviewer concluded that the performance of all groups was distinctively poor in physical science.

While statistics quoted out of context can be misleading, we do quote selected results in order to give some of the flavor of the first science test. The nine-year-olds knew a number of facts and principles: Iron cannot be burned in an ordinary fire (89 percent); day and night occur because the earth rotates (81 percent); there is a scientific reason why a rubbed balloon will adhere to a wall (78 percent). Given a beam balance, 96 percent were

* Material on the National Assessment can be obtained from: National Assessment of Educational Progress, 1860 Lincoln Street, Denver, Colorado 80203.

able to balance two weights on the beam and 94 percent were able to counterbalance two weights with a third. However, 69 percent responded that mixing equal quantities of water at 70 degrees and 50 degrees would result in an overall temperature of 120 degrees; only 7 percent gave the correct response.

The thirteen-year-olds knew that oxygen is the most abundant element in the human body (92 percent) and that lack of atmosphere on the moon precludes such activities as flying a kite or building a fire (74 percent). Half of them were aware of refraction as the phenomenon responsible for the apparent bending of a spoon at the surface of water; 61 percent identified the prime difference between hot and cold water as due to the speed of molecules. Only 4 percent were able to find the density of a wood block using the beam balance and the weight of a known mass; only 78 percent were able to time correctly how long it takes a pendulum to swing back and forth ten times. A majority (64 percent) stated that they are "sometimes" curious about why things in nature are the way they are, although only 8 percent indicated they "often" had such curiosity. Most (94 percent) believed that women can be successful scientists.

Many seventeen-year-olds (69 percent) knew that a galaxy contains many stars and that electric current in a copper wire involves mainly the movement of electrons. However, only 54 percent knew that the speed of a falling rock increases as it falls, 46 percent that the higher of two musical notes has the higher frequency and shorter wavelength, and 56 percent that mercury can be enclosed in glass to make a thermometer because it expands more than glass when both are heated together. Asked what metal cans for holding foodstuffs are chiefly made of, 93 percent said tin, while only 3 percent chose iron. Seventy-five percent used the beam balance correctly in an exercise requiring the use of weights and distances from a pivot point; 56 percent timed the pendulum correctly, but only 12 percent were able to determine the density of wood, whereas 63 percent gave an incorrect answer and 25 percent did not answer.

Adults scored best in questions that could be answered with information gained by reading newspaper and magazine articles or watching television. Sixty-four percent knew the function of a fuse in an electric circuit; 85 percent were aware that the movement and characteristics of air masses are important in predicting weather. Fifty-six percent were able to compute the time required for a boat to travel downstream, given the speeds of the boat and river. Half knew that a rock gains speed as it falls. The beam balance exercise was performed correctly by 74 percent of the adults, but only 12 percent could determine the density of wood. The swings of the pendulum were timed correctly by 49 percent.

This brief survey indicates that most of the questions were searching for factual knowledge, an emphasis inconsistent with the educational goals we have been espousing. This class of questions tends also to be answerable on the basis of information learned outside of school.

The success of the assessment will depend mainly on the extent to which each question is directed unambiguously toward the objective it is designed to test. The panelists expressed a number of doubts about the trueness of aim. A critical review of the questions by an outside panel would appear more valuable before the test is used, rather than after, with determination to make even extensive modification in response to valid objection.

The physics teaching community undoubtedly would have predicted the outcome of the assessment in physical science suggested by certain of the review panelists and also that this outcome will be validated by future, more reliable assessments. However, "I told you so" is not a fitting response, for this same community has long held the key to a better outcome.

11.2 THE FEDERAL PRESENCE IN SCIENCE EDUCATION

Since its beginning in 1950, the National Science Foundation has been the principal agent for conducting federal programs in science education. Its many programs and activities throughout this period can be grouped roughly into three categories: encouragement and assistance for students, strengthening the qualifications of teachers, and the development of improved courses and curricula. Most of this effort has been allocated to the education of professional scientists, but the NSF has assumed increasing responsibility for the improvement of science education at all levels and for a broader range of educational objectives. The resources given this agency by Congress and the Executive Branch have enabled it to have a large quantitative impact on the professional education of scientists and even to exert a significant qualitative effect on high school physics. However, its role in education more broadly must be limited to pilot efforts at innovation, for it does not have the resources to mount, for example, a major national program to improve elementary school education by reeducating a million teachers.

Although local school systems, state departments of education, and a number of federal agencies share responsibility for science education, the NSF is unique among them in its close link with the scientific community. When educational improvement can be regarded as an adjunct or an outgrowth of their other activities, scientists find it easier to participate. The inclusion of the improvement of science education as an explicit and strong function of the NSF is the best current assurance of accomplishing this participation.

Further conditions are necessary if the NSF's educational activity is to have a quality matching its importance. Because science education is part of science, much of its style and even its vocabulary are those of science. Communication between the education staff of the Foundation and the outside scientific community is largely scientific discourse. It would take place haltingly and uneasily were not the personnel of the Foundation scientists. But there is a further, more stringent condition. The programs formulated by the NSF and their implementation and evaluation are the responsibility of the staff of its Education Directorate. Obviously, the quality of these activities is dependent on the quality of the staff. The staff, however, does not act in isolation; it has continual interaction with the scientific community through advisory panels and committees, reviewers of proposals submitted to the NSF, and informal contacts at meetings, to mention only the most prevalent mechanisms for exchanges of views. The quality of the advice and suggestions depends on the caliber of the persons with whom the staff interacts, and the nature of the staff is a major determinant of high-quality, productive interaction with the scientific community.

One can state with justice, therefore, that the argument for locating federal programs for science education in the NSF depends heavily on the ease with which its personnel can draw on the assistance of the best of the nation's scientist–educators. The caliber of the staff is the key. Though mindful of the NSF's many successes, we believe complacency to be unwarranted and urge that a continuing effort be made to attract ever more able persons to the education staff.

The task of surveying the NSF's education program is made easier by a report issued in 1970 by its Advisory Committee for Science Education.[101] This report reviews the Foundation's educational programs, especially the course content improvement activities, evaluates them, and recommends that emphasis be shifted toward scientific education of the public. In general, and in most of its particulars, our report is consistent with this earlier one; it, too, recommends greater effort toward further precollege curriculum development, preservice teacher education, the improvement of courses, and the teaching of science for the nonscience undergraduate. It also advocates greater perspective on social needs in the education of scientists, as well as in the training of technical personnel, and reconsideration of the nature of advanced degrees. In some respects there are differing views: for example, the NSF report calls for greater emphasis on interdisciplinary, problem-oriented education for scientists than we do.

As the NSF begins to stress the task of improving science education for the nonscientist, it should not have to phase out programs of proven merit that help to maintain the strength of science in the United States. Predoctoral fellowships offered through national competition are in this category. The winning of an NSF Graduate Fellowship has become an undisputed mark of promise. A student holding one is welcome anywhere; he is able to develop as a scientist in the surroundings he thinks best suited to his purpose. Surely, society gains from providing this kind of assistance for its most promising scientists.

The U.S. Office of Education is entrusted with the expenditure of about $5 billion per year. In contrast, the NSF obligation for education reached its highest point, approximately $125 million, in 1968 and has declined to about $70 million per year since that time. The budgets of the two agencies reflect the difference between the specialized concerns of NSF and the quite different congressional and executive intent behind the appropriations for the Office of Education. The Office of Education has a broad mandate to support education at all levels, but most of its funds are distributed to the states and local school districts on a formula basis. It supports research and instruction in educational methodology. The improvement of subject matter instruction is a less direct concern. However, if new science curricula are to be disseminated widely, resources more on the scale of those available to the Office of Education will be needed.

We recognize a strong need for close cooperation between the Office of Education and the NSF if the objective of strengthening science education for the general public is ever to be achieved. Given the tested channels of communication between the science community and the NSF and the dearth of such channels in the Office of Education, it seems reasonable to suggest that most innovative and developmental efforts be conducted through the NSF, at least for the foreseeable future. Yet, to offer a substantial fraction of the in-service teachers improved preparatory study will require a better understanding of the learning process. In addition, to facilitate the distribution of science materials will require expenditures and channels to the schools that at present only the Office of Education is equipped to manage. To attack the problem by, for example, dividing the responsibility for science education according to the educational level, assigning the lower levels to one agency and the higher levels to the other, would present disadvantages, for this policy would tend to remove a part of science education from the province of the scientists and thereby weaken their interest and involvement. It is gratifying, therefore, to find in the 1970 Report of the NSF Advisory Committee on Education a call for strengthening the liaison between the two agencies.

The drastically diminished level of support for educational activities in the NSF has been a matter of great concern, although it is recognized that much of the decrease is a consequence of a deliberate reversal of a policy instituted by the federal government a decade ago to accelerate the production of scientists and engineers. Accompanied as it is by mandates to expand programs for the improvement of general science education and to emphasize the applicability of science to the achievement of socially desirable goals, this budgetary pressure is felt intensely by programs intended to improve the curricula and facilities for educating future scientists and engineers. Society cannot ignore the reality of the costs of education for science; like the study of medicine, it always has been relatively expensive. But the nation as a whole gains from the productivity of scientists and engineers and should not shirk the task of educating them effectively.

Legislation has been introduced to establish two new educational agencies, the National Institute of Education and the National Foundation for Higher Education. In the words of President Nixon in his education message to the Congress in March 1970, the National Institute of Education would be aimed at a "serious, systematic search for new knowledge needed to make the educational opportunity truly equal." It would be located in the Department of Health, Education, and Welfare. Again in the President's words, "It would have a National Advisory Council of distinguished scientists, educators and laymen to ensure that educational research in the Institute achieves a high level of sophistication, rigor and efficiency." No one who looks critically at the present output of educational research will question the need for a central means of coordinating and evaluating it in the way indicated by the President, whether it be directed explicitly at the goal of equalizing educational opportunity or at other goals. If carried out in a truly critical spirit, the evaluation of educational research could lead to the recasting of education to give all people a more realistic opportunity for satisfaction from education.

The National Foundation for Higher Education, which would be much like the NSF and the National Foundation for the Arts and Humanities, would have as its objectives the support of ideas for improvement and reform in higher education, the strengthening of colleges and universities that have special needs or that fulfill special functions, and the development of a national policy on higher education. To insure coordination, the Director of the NSF and the Commissioner of Education would serve *ex officio* on the Higher Education Board, which would govern the new Foundation.

Should one or both of these agencies come into being, the effect on education in the United States could be profound. In particular, of course, the education programs of the NSF would be affected.

11.3 ACCREDITATION?

Accreditation of physics departments is viewed by some as a guarantee that students motivated toward physics will be guided away from institutions offering poor or inadequate education in this subject. We believe that formal accreditation is undesirable under any circumstances and that there are more positive ways to accomplish not only the professionally oriented objective mentioned but also improvement of physics education everywhere.

Our reasons are the following: There are a variety of professional societies in physics and an even greater diversity of occupations for physicists; to attempt to formulate standards acceptable to all would be a difficult and probably fruitless process. Rather than being imposed by an external agency, standards should be a concern of the individual educational institution. Each institution has its own objectives, its own view of itself, and it has access to outside opinion when it so desires. Quality is difficult to define, and high-quality education, no easier. Listings of courses, laboratory apparatus, and facilities, even the citing of textbooks used, give only the facade of education; distinctions won by faculty members provide another dimension but are not likely to be prescribed by an accreditation team. A set of standards once established is difficult to change, and an institution once accredited may find little incentive for further improvement. Those who would innovate need encouragement rather than the reverse. Finally, the students in many accredited institutions would find their instructors performing in ways dictated from the outside rather than from inner convictions.

In the physics community, there is a broad consensus regarding those institutions that mount successful programs of instruction. We support the view that periodic publication of the curricula and practices of these institutions will provide a context in which all can view their own efforts. If made available to students, compilations of this nature can provide guidance in selecting appropriately oriented undergraduate and graduate programs. The AIP publishes a directory of graduate programs and in the early 1960's published information about selected but unidentified public and private undergraduate programs.[105,106] We recommend an AAPT program to prepare and publish periodically objective descriptions of the curricula and facilities available for the teaching of physics in all those institutions that offer an undergraduate major in this field. Admittedly a monumental undertaking, it may well require a continuing effort, as the task of completing a survey of 600-odd programs at any one time would require resources beyond those foreseeable.

11.4 PHYSICS AND THE DISADVANTAGED

The problems confronting the United States in its effort to offer racial minorities and other disadvantaged citizens constructive and liberating outlets for personal development touch all facets of life. Physics, which as a profession supports only a tiny fraction of the population directly, would appear at first sight to merit little consideration as a tool for achieving this national goal. The professional activities of physicists do not appear to lead to goods or opportunities having short-range value for this purpose. The identification and encouragement of scientifically talented young disadvantaged people is, of course, an obvious and desirable activity, but again, this does not touch the general problem. However, a growing number of educators, from both minority and majority groups, have seen ways in which physics can serve the broader purpose.

Morris R. Lerner, a high school teacher and president of the National Science Teachers Association, has stated the case in these terms[107]:

> What has physics to offer the black student from the ghetto? As the courses are presently constituted, little or nothing. But I believe that physics has the greatest potential for moving these students into the mainstream of American life. The way out of the ghetto is money, and to get money one must get a good job. These students know this and it is their primary goal. Here is where physics can play a great part. We have a fascinating subject with almost built-in motivation, and we can use it to help these students develop the linguistic and mathematical skills they lack, and at the same time learn basic principles and approaches to problem-solving which can be forever useful to them. *But not in our traditional manner.* In courses for these students the specific content chosen is not particularly important except that it must be relevant to the lives of the students and amenable to development in a laboratory setting in small steps. The purpose of such courses is to give the student opportunities to use the methods of science and to apply these methods to solving problems that are of interest and significance to him. For these students physics can be of value as a specific in a technological society, but perhaps its greatest value is as a device to achieve a larger educational end.

With many variations in detailed approach, programs with the objective just described are being tried in various regions and with various disadvantaged groups. Most of these projects are directed toward the high school level and are based on collaboration between the schools and universities or colleges. Many comprise summer programs; such, for example, is a project at the University of Colorado[108] whose object is to motivate Chicano high school students to go on to college. High school sophomores are selected, who, on completion of the summer program, participate in follow-up activities for two years.

Project Beacon,[109] in New York City, involves the cooperative efforts of the high schools, York College of the City University of New York, and local industry. Its intent is to stimulate interest in science and engineering among blacks and Spanish-speaking students and to improve the high school curriculum for them. It operates on two interrelated levels, those of students and teachers.

Drawing heavily on two sources of material treated earlier in this report, PSNS and Project Physics, Project Beacon takes a flexible approach toward learning, although experience has taught the need for careful structuring, at least at the beginning. An explicit aim of the summer phase of the project is "to demonstrate to the teachers that the inquiry method, combined with an emphasis on student laboratory work, is eminently suited for accomplishing the goals of student motivation and orientation." From this phase, "the teachers saw that students from disadvantaged backgrounds can be motivated towards science if they are permitted to handle equipment and if they are encouraged to work independently in a relaxed setting."

The academic-year phase was conducted during six periods per week at the high school and an additional two hours per week at York College. Industry participated by providing career-oriented summer jobs. Teachers in the project reported[110] on a number of problems they encountered and techniques they found useful in coping with these problems. Initially skeptical of the proposal to use Project Physics materials with this population, they now believe that the success of the program in reaching the disadvantaged student consists of a judicious combination of these materials with the standard curriculum, with particular care being taken to accommodate to the student's abilities and background.

In the belief that inner-city schools need teachers with broader education in the sciences than the average, that their teachers need to have a commitment to teach in this environment, together with an understanding of inner-city life and culture, and that inner-city schools, perhaps beyond all others, require the revitalization of academic excellence, a special science teacher training program is under way at Brown University.[111] The departments of physics, chemistry, biology, and education cooperate in this effort with the public school system of Providence. The program prepares students for the AB degree in science education, with an optional fifth year leading to the MA degree. Most of the courses have been developed specifically for the program. They stress the unity of the sciences, emphasize student participation rather than formal lectures, discuss the role of science and technology in modern society, familiarize the student with traditional and new secondary school curricular materials, and provide direct interaction with high school students and teachers. Advanced courses are offered to enable the student to specialize in one of the disciplines.

The study of physical sciences, far more than the humanities and social sciences, appears *a priori* to be "color blind." For this reason, there have been no strident demands for racially relevant physics courses. However, precisely because of its freedom from emotionally charged issues, science study can be a useful vehicle for bringing about constructive social change.

Programs of recent AAPT meetings give evidence that an increasing number of physics instructors are beginning to use their subject, without demeaning it, in a purposeful way. Today's much-touted quality of relevance is an inherent property of their efforts.

11.5 PHYSICS IN THE BLACK COLLEGES

Higher education for those from disadvantaged backgrounds has a special feature in the United States, the existence of the many black colleges. Roughly half of all black undergraduates attend these colleges. Black colleges therefore are a significant national resource that must be used effectively if the nation is to achieve its goal of educational opportunity for all its citizens.

In 1968, the question of college physics for minority group students was explicitly raised in the Commission on College Physics and in the American Association of Physics Teachers, with the result that a symposium on Physics and the Nation's Racial Problems was held at the June 1969 AAPT meetings and a joint AAPT-CCP committee was formed to continue to focus attention on the problems.

In January 1970, this committee, with AIP sponsorship, convened a conference of the physics faculties in black colleges and other institutions with large enrollments of minority group students. A report of this conference, "Physics in the Black Colleges," is available.* The conference characterized the general problems for physics education that have been created by the environment surrounding most blacks in terms of three major components: (a) Because of their environment, many blacks have lessened confidence in themselves and their schools, (b) many pressures draw students away from physics, and (c) many students have deficiencies in skills essential to physics. As a consequence, the needs of entering black college students overwhelm the resources of the colleges, and the instructors must spend enormous amounts of time providing remedial instruction and inspiration and initiative to their students. Moreover, a theme that has appeared earlier in this report is equally pertinent: Introductory physics is seldom gauged to the needs and experiences of students. In a counterproductive way, the desire for accreditation of the college as a whole, which usually requires a minimum fraction of PhD-holding faculty, forces these schools into expensive competition for research-oriented faculty. While the physics major in a black college is almost guaranteed small classes, he

* Copies are available from the Department of Physics, Morehouse College, Atlanta, Georgia.

is likely to feel isolated from other students and to lose motivation. Physical facilities for him are now either sufficient or are rapidly becoming adequate.

The conference saw an urgent need to increase the professional caliber of physicists in black high schools and colleges and to develop curricula along lines already pointed out. To implement the drive for improvement, it called for the creation of regional centers devoted to research and teaching, summer workshops, and a clearinghouse for information relevent to the teaching of physics to minority group students, all of which might be included in the developments discussed in Chapter 7. Finally, it pointed out that, to a first approximation, all black students need financial assistance.

11.6 PHYSICS OUTSIDE THE CLASSROOM

Today's world, with its bewildering array of materials both natural and synthetic, its sophisticated manufacturing techniques, its wealth of display media, and its mobility, offers rich opportunity to learn in less formal contexts than the classroom. There are many motives for exploiting these possibilities. They provide ways for the public to keep in touch with discovery and new developments; a complementary aspect is the occasion they offer the scientist for describing his work. They provide additional stimuli for pupils in the schools. They can release science from its usually serious mood and show it to people in a lighter—even a playful—one.

Television is an obvious key to informal public education. In some parts of the world, it is also an effective means of offering science information to the public. The British Broadcasting Corporation's second channel presents an hour-long scientific documentary program each week, 40 weeks a year (one director, Alec Nisbett, is a theoretical physicist; several others have similar scientific backgrounds) that is directed at an audience similar in level of information and education to the readership of the *Scientific American.* The audience has reached 5 million, or 11 percent of the nation. The competition faced by this program is greater than that met in many parts of the United States, excluding the larger cities; the commercial and highly competitive BBC-1 channel schedules some of the leading drama and comedy series against it. But in comparison to the United States, the response to scientific entertainment in the United Kingdom is immense. It is generally believed in the television industry that such a program would fail in the United States. Yet, recently in this country, two other British productions, one in science and one on culture more generally, have been shown, with considerable success. The 2½-hour presentation, "The Violent Universe," on recent discoveries in astrophysics and cosmology, was shown on National Educational Television (NET) and later by a major commercial network. A commercial firm underwrote the showing on NET of the series "Civilisation," made by the BBC and Kenneth Clark. The time is ripe to apply existing expertise at the level of support and competence used by the BBC in these examples to produce a series on all science, with appropriate emphasis on physics and astronomy, for wide distribution in the United States. It should cover science at the level of the educated layman. The American Institute of Physics may be an appropriate agency to initiate and coordinate this effort.

In one western city of the United States, there is an Exploratorium, the manifestation of a physicist's idea for inviting people to learn for themselves. Some would call it a science museum, but its originator dislikes the connotations of this term. The notion is to open a place for activity and have materials and tools available, as well as many examples of scientific and technological processes, equipment, and demonstration experiments. Schoolchildren, college students, and passers-by come in and pick up objects that interest them, finding out how they work or constructing something from materials kept at hand there. In another city, there is a Science Center, frequented by schoolchildren, for whom it is intended, that also has a demonstrated fascination for adults. In this center is a "math" room, perpetually crowded with children drawing Lissajou's figures with a pendulum, making the Random Walker obey the directions of dice, creating soap film surfaces, playing mathematical games, and, generally speaking, participating rather than merely observing.

A number of cities have well-established, centralized science teaching resources. Nothing is new about these ideas, but, as with the examples cited, they can be modified to become creative responses to the contemporary scene. It is hardly necessary, much less possible, for many physicists to give the full effort required to initiate and carry through the development of an Exploratorium, although the field is far from saturated. However, the successful components, or even whole operations, can be copied with only moderate outlays of a scientist's time. Further, physicists can support and contribute to the spread of this kind of education at a variety of levels of commitment.

It is now easy to have one's own Exploratorium at home. Toys, notions, and ornaments explicitly illustrating physical phenomena have always enjoyed a certain amount of popularity among adults as well as children and continue to do so, even in the "Age of Aquarius." In terms of economics, there would seem to be considerable leverage; of the \$2.3 billion toy market in 1964, it was found that 5 percent comprised educational and scientifically based toys.[112] The extent to which physicists have contributed to the ideas for toys is not known, but many toys are recognizable as variants of devices commonly used in physics classes. Because the objective of a toy is to elicit free response, physicists should welcome those that generate spontaneity and suggest questions, in contrast to those that lay out a structured program and thereby mimic the classroom.

A rich variety of materials is readily available today, and at very low cost, that make play of a kind unthinkable a few years ago quite inexpensive now. Hand-held, manually operated vacuum pumps, lenses, and the like—catalogs abound with such objects. Elizabeth Wood has discovered how to use the television set as a stroboscope to show surface tension ripples on the surface of water in a pie tin. Her small book, *Science for the Airplane Passenger,*[113] is written in this vein. To design a program for the exploitation of today's commonplace materials is not the task of this report, but to point out a happy way of providing informal contact with phenomena and concepts that aid in developing what is called "physical insight," cannot be amiss.

11.7 LEARNING AIDS

The learning process can be aided in many ways, including improvement in the format, or system, of course organization or the introduction of useful items of equipment or printed works. In this section, we comment on one currently prominent example of each of these two means.

In Chapters 4, 5, and 6, we have discussed the validity of adapting the pace of instruction to the pace of learning. For some students, the conventional system of lecture, laboratory, problem set, and examination is satisfactory, but for many it is not. During the past five or six years, increasing attention has been given to an alternative system, commonly called the "Keller Plan" after the psychologist, F. S. Keller, the senior member of a team that first introduced a working model of a self-paced system[114] that has been tried by a number of physics instructors.[115]

The Keller Plan is a form of what is variously called "self-paced" or "personalized" study. It is based on the division of the coursework into units, in some examples consisting of about one week's work of a conventional course (or approximately the ten-hour unit of Chapter 3). The instructor prepares in advance a study guide and a set of about four tests for each unit. The study guide tells the student what he is supposed to learn from the unit and offers reading suggestions, problem sets, and explanatory material needed to supplement the textbooks. When the student feels prepared, he requests a test, writes out his answers, and brings his paper to a tutor (who is usually a well-prepared undergraduate who receives academic credit for assisting with the course). The tutor reads the test paper on the spot and discusses it with the student. If the student passes, he is given the study guide for the next unit; if not, he is directed to study further and to apply again for a test. The tutor is responsible for about ten students. In addition to the tutor, teaching assistants are assigned for laboratory work (some units might be completely lab-oriented). The faculty instructor who carries the principal responsibility is a court of appeal. He also gives lectures, but in most manifestations of the plan, only a few lectures per term. These are open only to students who have completed a specified set of units and after they are on subjects tangential to the prerequisite units. Grading has been accomplished in a variety of ways. In some instances, midterm and final examinations are given; in others, grades depend on the number of units completed (each unit being graded on a pass–fail basis).

Green's article gives considerable detail about MIT's version, discussing its variations, limitations, and cost and its impact on students, which seems to be felt primarily because of its self-paced feature, the freedom to study where and when one wishes, and the frequent contact with the tutor. Touted by no one as a panacea, and incorporating a number of familiar elements, the Keller Plan has caught the attention of a considerable number of college and university instructors as a practical departure from traditional methods that appeals to the ethos of the times.

That scientists should initiate the application of computers to instruction is only reasonable. Some of the earliest efforts to do so were made in the 1960's; one of the first concerted efforts in writing interactive programs in physics occurred at a conference on new instructional materials held by the Commission on College Physics in 1965. During that same year, an additional CCP conference was devoted entirely to computer use in undergraduate instruction. *The Computer in Physics Instruction*[94] was a result distributed to thousands of educators and scientists. During 1970, a second major convocation was initiated by CCP but with broader coverage: The Conference on Computers in Undergraduate Science Education, Physics and Mathematics.[97] A conclusion from the conference is that no general procedure yet exists for the widespread use of computers in education. Many had hoped, and even predicted, that by this time computers would have found widespread use, but the conference showed that this goal is not likely to be achieved soon. In various areas of special application, such as the computational mode, computer graphics, the simulation mode, analogue computing, and computer-supervised instruction, progress has been made, and on occasion real success, such as the PLATO project, is scored.

Inevitably, a younger generation accepts and uses new techniques, while an older one expresses its doubts. The automobile was cited in the 1970 conference by A. G. Oettinger,[97] as having started as a horseless carriage and eventually having transformed society. Not all this transformation is now seen as desirable, or even healthy, but society will unquestionably accept some of it. Oettinger (and the conference in general) feels the same thrust of inevitability in regard to computers in education. Oettinger believes that, because of the greater flexibility of universities, which facilitates innovation and experimentation, the development and use of computers for education will grow in these institutions before computer-aided instruction becomes widespread in the schools. The federal agencies view computers as a possible means of meeting the crisis that the simultaneously rising costs and standards for education could bring about. Both the NSF and the Office of Education are supporting research and development projects to further the application of computers to education. Physics teaching can play a major part in this great national effort.

11.8 INTERNATIONAL ACTIVITIES

Physics education presents problems with a high degree of international commonality despite strong differences in the educational frameworks adopted by the many nations. A comprehensive guide to the variability of physics education can be found in a comparative study, *A Survey of the Teaching of Physics in Universities*,[116] which also describes preuniversity education. Here one will find diagram and description, examination question and laboratory practice, if he wishes to understand the systems that have produced the colleagues he meets when he attends a conference abroad or an international gathering at home.

The improvement of physics teaching has been explored in a sequence of international conferences sponsored by the International Commission on Physics Education of the International Union of Pure and Applied Physics. These conferences have considered many of the matters discussed in this report, including general education in physics, high school physics, the curricula for undergraduates and graduates, and the preparation of secondary school teachers of physics.[117-120] The International Commission has sponsored two seminars as well, one on the Education of Physicists for Work in Industry and another on the Role of the History of Physics in Physics Education.

In addition to the *Survey* mentioned above, UNESCO and the International Commission on Physics Education have been engaged jointly in publishing a series on *New Trends in the Teaching of Physics* and a *Source Book for the Teaching of Physics in Secondary Schools.*

UNESCO also has undertaken projects in science education in a number of countries—in Asia, Africa, the Near East, and Latin America—that need special help. The last of these included a one-year pilot project on the teaching of physics followed by a series of regional conferences for physics teachers. In many of these areas, permanent organizations have been established to carry on developmental work.

A recent review of these activities[120] expresses the main worry of UNESCO as "the limited participation of physicists and physics teachers from the developing countries," and addresses the question of how to identify and help those with potential for becoming leaders in science education in these countries. Organizations that support promising young scientists from developing nations when they study in the more advanced countries, and the institutions that receive these young people, are asked to pay more attention to this need as they devise programs to accommodate them.

11.9 OPEN UNIVERSITIES

We view with considerable interest the Open University, which began its teaching programs in England in 1971.[121] It is intended primarily for adult students employed full time or working in their homes. No formal academic qualifications are required. It is "not simply an educational rescue mission" according to its Chancellor, nor is it a rival of the existing universities. It offers six main lines of study: arts, science, mathematics, technology, social sciences, and educational studies. It awards a BA degree and the degrees of Bachelor of Philosophy, Master of Philosophy, and Doctor of Philosophy as well as Doctor of Letters and Doctor of Science. It operates as a correspondence school with radio or television lectures. Packets of course materials are mailed to students, and the lectures are given at weekly intervals. Students have tutors and examinations. Two-week summer schools are mandatory for some courses. There are plans to establish study centers in locations in which the density of students is sufficient. A recent report on the Open University states that students in a second-level course in electromagnetics and electronics receive packets containing an oscilloscope, a signal generator, a dc supply, and other components. It also relates that the attrition rate for students entering the second year is only half of that originally expected, so the new first-year intake has been reduced by 20 percent to 20,000 of the 35,000 applicants. Many overseas institutions are purchasing supplies from the Open University to initiate their own versions.

A recent article[122] provides evidence that in a number of regions of the United States educators have seen in open universities an opportunity to meet the desires of a much more varied component of the population than do the traditional educational structures. Among the ventures planned are national universities, offering enrollment to students anywhere in the nation. One of these would act as a distributor of the British materials, with credit to be granted by individual colleges and universities. Another would grant credit for on-the-job training, internships, courses at local colleges, and courses offered by institutions other than colleges. Degrees would be granted on the recommendation of a council of academic advisors from various fields of study. A number of the ventures are conceived as adjuncts to existing state systems of education, and one is the creation of a group of public and private colleges and universities scattered over the East and Midwest.

Applications for admission to these new programs still are small in number, and, since most of them are in formative stages, enrollments are measured in the hundreds. But the apparent success achieved in Great Britain is a strong indication that this mode has great appeal and that their student bodies will grow rapidly.

The physics community may fully expect to be confronted with the need to prepare physics courses of a more self-contained nature than any existing ones. The vital phenomenological component will have to be provided; physicists will have to judge the success of the British venture in mailing kits. In addition, they almost certainly will have to establish well-equipped centers or open existing institutions for laboratory study that cannot be carried out in the home for one reason or another.

References

1. *Toward a National Science Policy,* Report of the Subcommittee on Science, Research and Development to the Committee on Science and Astronautics, U.S. House of Representatives, 91st Congress, Oct. 15, 1970.
2. M. Green, *Science and the Shabby Curate of Poetry* (W. W. Norton, New York, 1965), p. 74.
3. J. Priestley, *The History and Present State of Electricity, with Original Experiments* (J. Dogsley, J. Johnson, B. Davenport, and D. Codell, London, 1767), from the Preface.
4. W. R. Snelling and R. Boruch, "Factors Influencing Student Choice of College and Course of Study," J. Chem. Educ. *47,* 326 (1970).
5. J. R. Butler, *Occupational Choice* (Her Majesty's Stationery Office, London, 1968).

6. J. W. McKinnon and J. W. Renner, "Are Colleges Concerned with Intellectual Development?" Am. J. Phys. *39,* 1047 (1971).
7. J. Piaget, "Cognitive Development in Children," J. Res. Sci. Teaching *2,* 176 (1964).
8. J. W. Renner and D. D. Stafford, "Inquiry, Children and Teachers," Sci. Teacher *37,* 55 (1970).
9. P. E. Blackwood, *Science Teaching in the Elementary Schools* (U.S. Dept. of Health, Education, and Welfare, Office of Education, Washington, D.C., 1965).
10. U.S. Dept. of Health, Education, and Welfare, Office of Education, National Center for Educational Statistics, *Digest of Educational Statistics, 1970* (U.S. Govt. Printing Office, Washington, D.C., 1970).
11. The Physical Science Study Committee and Educational Services, Inc., *Physics,* 3rd ed. (Heath, Boston, Mass., 1970).
12. Harvard University, *The Project Physics Course* (Holt, Rinehart and Winston, New York, 1971).
13. Engineering Concepts Curriculum Project, Polytechnic Institute of Brooklyn, *The Man-Made World* (McGraw-Hill, New York, 1971).
14. "Enrollment Increase in Science and Mathematics in Public Secondary Schools, 1948–49 to 1969–70," *Science Resources Studies Highlights,* NSF 71-30 (National Science Foundation, Washington, D.C., 1971).
15. E. A. Wood, *Pressing Needs in School Science* (American Institute of Physics, New York, 1969).
16. International Conference on Education in Chemistry, Division of Chemical Education, American Chemical Society, *Preliminary Report* (ACS, Washington, D.C., 1970).
17. PSNS Project, Rensselaer Polytechnic Institute, *An Approach to Physical Science* (Wiley, New York, 1969).
18. C. E. Swartz, "Let's Do Away with High School Courses in Earth Science, Biology, Chemistry and Physics," Phys. Teacher *38,* 65 (Feb. 1971).
19. L. Deall, "Development and Implementation of a Modular Four-Year Secondary School Physics Program," Am. J. Phys. *37,* 1183 (1969).
20. American Institute of Physics, Division of Education and Manpower, *Attrition of Physics Students* (AIP, New York, 1969).
21. U.S. Dept. of Health, Education, and Welfare, Office of Education, National Center for Educational Statistics, *Projections of Educational Statistics to 1979–80* (U.S. Govt. Printing Office, Washington, D.C., 1971), p. 82.
22. American Institute of Physics, *Physics Manpower and Educational Statistics* (AIP, New York, 1962).
23. National Science Foundation, *Secondary School Science Teachers* (NSF, Washington, D.C., 1969).
24. National Science Foundation, *Secondary School Science and Mathematics Teachers* (NSF, Washington, D.C., 1963).
25. National Science Foundation, *Annual Report, 1970* (NSF, Washington, D.C., 1970). [It is stated here (p. 68) that nearly 50,000 individuals were expected to participate in institute programs during 1970–1971.]
26. W. Welch and H. I. Walberg, "An Evaluation of Summer Institute Programs for Physics Teachers," J. Res. Sci. Teaching *5,* 105 (1967).
27. S. G. Bronevshuk in *Teaching Physics–An Insoluble Task?* S. C. Brown, F. J. Kedves, and E. J. Wenham, eds. (MIT Press, Cambridge, Mass., 1971), p. 124.

28. F. Watson in *Teaching Physics–An Insoluble Task?* S. C. Brown, F. J. Kedves, and E. J. Wenham, eds. (MIT Press, Cambridge, Mass., 1971), p. 11.
29. H. E. Gruber, "Science as Doctrine or Thought? A Critical Study of Nine Academic-Year Institutes," J. Res. Sci. Teaching *1*, 124 (1963).
30. E. P. Torrance, *Education and Creative Potential* (University of Minnesota, Minneapolis, Minn., 1965).
31. J. H. Wilson and J. W. Renner, "The 'New' Science and the Rational Powers: A Research Study," J. Res. Sci. Teaching *6*, 303 (1969).
32. A. B. Arons, "Teaching the Sciences" in *The Challenges of College Teaching*, S. Cahn, ed. (Dr. Arons kindly supplied a copy of this work in advance of publication.)
33. Commission on College Physics, *Preparing High School Physics Teacher II* (CCP, University of Maryland, College Park, Md., 1972).
34. R. N. Little, "Pre-Service Formation of Physics Teachers: Technical Education," in *Teaching Physics–An Insoluble Task?* S. C. Brown, F. J. Kedves, and E. J. Wenham, eds. (MIT Press, Cambridge, Mass., 1971), p. 58.
35. J. D. Herron, Sci. Educ. News (AAAS) (July 1971).
36. A. A. Strassenberg, "Physics Teachers," Sci. Educ. News (AAAS) (July 1971).
37. H. R. Crane, "Students Do Not Think Physics is 'Relevant.' What Can We Do About It?" Am. J. Phys. *36*, 1137 (1968).
38. D. L. Abell, "A Working Conference on Source Material in Physics, Biology, Agriculture and Natural Resources," CUEBS Working Paper no. 2 (Commission on Undergraduate Education in the Biological Sciences, Washington, D.C., 1970).
39. W. H. White, *A Complete Physics. Written for London Medical Students and General Use* (Richard Clay and Sons, London).
40. R. G. Page and M. H. Littlemeyer, eds., *Preparation for the Study of Medicine* (University of Chicago Press, Chicago, Ill., 1969), Table 8.
41. P. Morrison, "Less May Be More," Am J. Phys. *32*, 441 (1964).
42. B. G. Aldridge, "Physics in Two-Year Technical Curricula," Phys. Teacher *8*, 302 (1970).
43. U.S. Dept. of Labor, Bur. of Labor Statistics, *Technician Manpower, 1966–1980* (U.S. Govt. Printing Office, Washington, D.C., 1970).
44. *The Junior College and Education in the Sciences*, Report of the National Science Foundation to the Subcommittee on Science, Research and Development of the Committee on Science and Astronautics, U.S. House of Representatives, 90th Congress, 1st Session (1967), p. 101.
45. U.S. Dept. of Health, Education, and Welfare, Office of Education, *Criteria for Technician Education–A Suggested Guide*, USOE 80056 (Washington, D.C., 1968).
46. *The Junior College and Education in the Sciences*, Report of the National Science Foundation to the Subcommittee on Science, Research and Development of the Committee on Science and Astronautics, U.S. House of Representatives, 90th Congress, 1st Session (1967).
47. National Science Foundation, *Junior College Teachers of Science, Engineering, and Technology, 1967* (NSF, Washington, D.C., 1967).
48. Commission on College Physics, *Preparing Two-Year College Physics Teachers, A Workshop Report* (CCP, University of Maryland, College Park, Md., 1972).
49. M. R. Parlett and J. G. King, *Concentrated Study*, Research in Higher Education Monograph No. 14 (Society for Research into Higher Education, Ltd., London, 1971).

50. R. P. Feynman, R. B. Leighton, and M. Sands, *Lectures on Physics*, 3 vols. (Addison-Wesley, Reading, Mass., 1965).
51. A. C. Helmholz and E. M. Purcell, *The Berkeley Physics Course*, 5 vols. (Mcgraw-Hill, New York, 1967).
52. J. I. Shonle, "A Report on Open End Laboratories," Am. J. Phys. *38*, 450 (1970).
53. American Institute of Physics, *Graduate Programs in Physics, Astronomy, and Related Fields* (AIP, New York, 1971).
54. Carnegie Commission on Higher Education, *Less Time, More Options* (McGraw-Hill, New York, 1971).
55. S. D. Ellis, *Work Complex Study* (American Institute of Physics, New York, 1969).
56. G. Diemer and I. H. Einch, *The Education of Physicists for Work in Industry*, The Proceedings of an International Seminar on Education of Physicists for Work in Industry (Centrex, Eindhoven, The Netherlands, 1968).
57. CCP Newsletter No. 25 (Nov. 1971).
58. J. M. Fowler, "Progress Report of the Commission on College Physics for 1966–1968," Am. J. Phys. *36*, 1035 (1968).
59. W. V. Johnson, "Pacific Northwest Association for College Physics, A Many-Body Problem," Am. J. Phys. *38*, 741 (1970).
60. J. D. Spangler and C. E. Hathaway, "The Consortium for the Advancement of Physics Education," Am. J. Phys. *39*, 399 (1971).
61. P. H. Coombs, *The World Educational Crisis* (Oxford U.P., New York, 1968).
62. C. E. Silberman, *Crisis in the Classroom* (Random House, New York, 1970).
63. N. Sanford, *Where Colleges Fail* (Jossey-Bass, San Francisco, Calif., 1968).
64. E. K. Faltermayer, "Let's Break the Go-to-College Lockstep," Fortune *LXXXII* (5), 98 (Nov. 1970).
65. I. Berg, *Education and Jobs: The Great Training Robbery* (Beacon, Boston, Mass., 1971).
66. A. E. Eurich, ed., *Campus 1980* (Delta, Montreal, 1968).
67. American Council on Education, *Universal Higher Education: Costs and Benefits* (ACE, Washington, D.C., 1971).
68. E. A. Dunham, *Colleges of the Forgotten Americans* (McGraw-Hill, New York, 1969).
69. P. Woodring, *The Higher Learning in America: A Reassessment* (McGraw-Hill, New York, 1968).
70. C. Jenks and D. Riesman, *The Academic Revolution* (Doubleday, Garden City, N.Y., 1968).
71. A. M. Heiss, *Challenges to Graduate Schools* (Jossey-Bass, San Francisco, Calif., 1970).
72. I. Illich, *Deschooling Society* (Harper and Row, New York, 1971).
73. A. Toffler, *Future Shock* (Random House, New York, 1971).
74. *Report of the White House Conference on Youth* (U.S. Govt. Printing Office, Washington, D.C., 1971).
75. T. Roszak, *The Making of a Counter Culture* (Anchor, Garden City, N.Y., 1969).
76. K. Keniston, *The Uncommitted* (Dell, New York, 1965).
77. K. Keniston, *Youth and Dissent* (Harcourt Brace Jovanovich, New York, 1971).
78. Carnegie Commission on Higher Education, *The More Effective Use of Resources: An Imperative for Higher Education* (McGraw-Hill, New York, 1972).
79. M. Blang, ed., *Economics of Education 2* (Penguin, Baltimore, Md., 1969).

80. K. Baier and N. Rescher, eds., *Values and the Future* (The Free Press, Glencoe, Ill., 1971).
81. C. W. Churchman, *The Systems Approach* (Delta, Montreal, 1970).
82. J. W. Gardner, *Self Renewal* (Harper and Row, New York, 1965).
83. President's Science Advisory Committee, Panel on Educational Research and Development, *Innovation and Experiment in Education* (Washington, D.C., 1964).
84. Science *173*, 1109 (Sept. 17, 1971).
85. U.S. Congress, House of Representatives, Authorizing Appropriations to the National Science Foundation, H.R. 92-977, p. 96.
86. P. J. Smith, "Britain's Open University: Everyman's Classroom," Saturday Rev. (Apr. 29, 1972), p. 50.
87. F. Newman, *Report on Higher Education* (U.S. Govt. Printing Office, Washington, D.C., 1971), p. 28.
88. R. E. Callahan, *Education and the Cult of Efficiency* (University of Chicago Press, Chicago, Ill., 1962).
89. P. E. Haggerty, "A Shift to the Production of Services," Science *173*, 679 (Aug. 20, 1971).
90. V. B. Smith, "More for Less: Higher Education's New Priority," in *Universal Higher Education: Costs and Benefits* (American Council on Education, Washington, D.C., 1971).
91. J. A. Gallant and J. W. Prothero, "Weight-Watching at the University: The Consequences of Growth," Science *175*, 381 (Jan. 26, 1972).
92. J. Rosenthal, "Performance Contracts Lost Luster," *The New York Times*, Jan. 10, 1972, Sec. E, p. 25.
93. L. P. Grayson, "Costs, Benefits, Effectiveness: Challenge to Educational Technology," Science *175*, 1216 (Mar. 17, 1972).
94. Commission on College Physics, *The Computer in Physics Instruction: Report of the Conference on Uses of the Computer in Undergraduate Physics Instruction* (CCP, University of Maryland, College Park, Md., 1965).
95. President's Science Advisory Committee, *Computers in Higher Education* (U.S. Govt. Printing Office, Washington, D.C., 1967).
96. A. G. Oettinger with S. Marks, *Run, Computer, Run: The Mythology of Educational Innovation* (Collier Books Edition, New York, 1971).
97. Commission on College Physics, *Computers in Undergraduate Science Education: Conference Proceedings* (CCP, University of Maryland, College Park, Md., 1971).
98. P. J. Esseff, "Comparative Costs of New Methods of Science Instruction," background paper for Commission on College Physics Conference on Priorities for Undergraduate Physics Education.
99. W. G. Bowen, *The Economics of the Major Private Universities* (Carnegie Commission on the Future of Higher Education, 1968).
100. "Sophistication Factor," Phys. Bull. *23*, 138 (1972).
101. Advisory Committee for Science Education, National Science Foundation, *Science Education–The Task Ahead for the National Science Foundation* (U.S. Govt. Printing Office, Washington, D.C., 1970).
102. L. F. Chin, "A Survey of Science Teaching in the Public Secondary Schools of the Great Lakes and Far West Regions of the United States in the 1970–1971 School Year," PhD Dissertation, Ohio State University (1971).

103. J. W. Maben, "A Survey of Science Teaching in the Public Elementary Schools of Two Selected Regions of the United States during the 1970–71 School Year," PhD Dissertation, Ohio State University (1971).
104. Education Commission of the States, *National Assessment of Educational Progress, Report 1, 1969–70 Science: National Results and Illustrations of Group Comparisons* (Denver, 1970).
105. American Institute of Physics, *Toward Excellence in Physics–Reports from Five Colleges* (AIP, New York, 1964).
106. American Institute of Physics, *Toward Excellence in Physics–Reports from Four Public Colleges and Universities* (AIP, New York, 1965).
107. M. R. Lerner, "Physics and the Nation's Racial Problems," Am. J. Phys. *38,* 126 (1970).
108. W. R. Chappell, "A College Motivation Program for Minority High School Students," AAPT Announcer *1,* 16 (1972).
109. F. R. Pomilla and M. S. Spergel, "The Model for Project Beacon," Phys. Teacher *9,* 136 (1971).
110. L. Siegel and R. Weinstein, "The High School Experience in Project Beacon," Phys. Teacher *9,* 134 (1971).
111. W. G. Massey, "Training Science Teachers for the Inner City: A High School and University Cooperative Program," AAPT Announcer *1,* 13 (1972).
112. A. J. Wood Research Corp., *Toy Buying in America: A One-Year Study Prepared for the Toy Manufacturers of America, Inc.* (Philadelphia, Pa., 1965).
113. E. A. Wood, *Science for the Airplane Passenger* (Houghton Mifflin, Boston, Mass., 1968).
114. F. S. Keller, "Good-bye, Teacher," J. Appl. Behav. Anal. *1,* 79 (1968).
115. B. A. Green, "Physics Teaching by the Keller Plan of MIT," Am. J. Phys. *39,* 764 (1971).
116. International Union of Pure and Applied Physics/United Nations Educational, Scientific, and Cultural Organization, *A Survey of the Teaching of Physics in the Universities* (G. J. Tieme N.V., Koninglijke Druppery, The Netherlands, 1966).
117. S. C. Brown and N. Clarke, *International Education in Physics* (MIT Press, Cambridge, Mass., 1960).
118. S. C. Brown, N. Clarke, and J. Tiomno, *Why Teach Physics?* (MIT Press, Cambridge, Mass., 1963).
119. S. C. Brown and N. Clarke, *The Education of a Physicist* (MIT Press, Cambridge, Mass., 1966).
120. S. C. Brown, F. J. Kedves, and E. J. Wenham, eds., *Teaching Physics–An Insoluble Task?* (MIT Press, Cambridge, Mass., 1971).
121. The Open University, *The Open University, Prospectus 1971* (Bletchley, Bucks., England, 1970).
122. P. W. Semas, "Open University Programs Gain Favor in U.S.," Chron, Higher Educ. *VI*(10), 1 (1971).

XIV
Dissemination and Use of the Information of Physics

PANEL MEMBERS

CONYERS HERRING, Bell Telephone Laboratories, Inc., *Chairman*
THOMAS LAURITSEN, California Institute of Technology, *Vice Chairman* (until May 1970)
FAY AJZENBERG-SELOVE, University of Pennsylvania
LEE GRODZINS, Massachusetts Institute of Technology (from January 1971)
ROBERT HERMAN, General Motors Corporation
GEOFFREY KELLER, The Ohio State University
ROBERT W. KEYES, International Business Machines Corporation
PHILIP M. MORSE, Massachusetts Institute of Technology
ARTHUR H. ROSENFELD, Lawrence Berkeley Laboratory
LOUIS B. SLICHTER, University of California, Los Angeles

Ex officio

LEWIS SLACK, American Institute of Physics

STAFF SCIENTIST

DALE T. TEANEY, International Business Machines Corporation

LIAISON REPRESENTATIVES

WALTER C. CHRISTENSEN, Department of Defense
WAYNE R. GRUNER, National Science Foundation
WALTER E. HUGHES, Atomic Energy Commission
A. W. SCHARDT, National Aeronautics and Space Administration

Contents

XIV
DISSEMINATION AND USE OF THE INFORMATION OF PHYSICS

1 Introduction: The Scope and Importance of Communication Activities

Communication of knowledge from one person to another is central to both the existence and progress of science and its economic and cultural impact on society. As a bare statement, this may seem trivially obvious; yet the full impact of the central role of communication is rarely appreciated. In this report, we hope to develop a picture of the present pattern of communication of physicists with each other and with the rest of the world—other scientists, technologists, students, and the public—and, after identifying some shortcomings in this pattern, to suggest at least a few possibilities for improvement. Many of the problems we shall address have already been discussed, in the broader framework of science and technology in general, in the Weinberg[1] and SATCOM[2] reports, whose wisdom and perspective we commend to the reader. However, limiting our discussion to physics will enable us to treat all the issues with much more concrete detail than was possible in these very general studies.

The value of efficiency in communication—getting the right information to the right people at the right time and without excessive consumption of the time of these people—is perhaps most obvious in education and in the relation of physics to technology. In both these cases, a translation is needed to adapt the lore of physics to a particular group of users. But communication is also the lifeblood of physics itself, as it is of any science. The intrinsically social nature of science has been convincingly set forth by the

British physicist J. M. Ziman in his book *Public Knowledge*.[3] His thesis, to which we subscribe, is that "The absolute need to communicate one's findings, and to make them acceptable to other people, determines their intellectual form. Objectivity and logical rationality, the supreme characteristics of the Scientific Attitude, are meaningless for the isolated individual; they imply a strong social context, and the sharing of experience and opinion." Consensus, without which science cannot go forward, is made possible only by communication and discussion in the public record. To quote Ziman once more, "Technology, Art, and Religion are perhaps possible for Robinson Crusoe, but Law and Science are not."

1.1 SCALE OF VALUE

Granted, then, that communication is the lifeblood of physics, how can we measure its importance, even within physics, in economic terms? In other words, how can we make even a rough estimate of the value of an improvement of some aspect of communication relative to the value of a given amount of research support? Ideally, one would like to do an experiment comparing the progress made by two matched groups of physicists, working on identical problems, one with the improved communication media and services and one without. But real life hardly ever provides such examples: Although it is easy to note that physicists in very small institutions or in remote countries are less productive than those in flourishing centers, there are many factors other than communication that might share responsibility for this lower productivity. There are, however, several less direct ways to get rough quantitative measures of value, four of which we shall now enumerate. The first two of them rely on the working judgments either of individual scientists or of small organizational groupings that are directly influenced by the thinking of many individuals; these judgments, though surely imperfect, are probably our safest source of information, since they are constantly being made and revised in response to pressures of work. The last two again involve grassroots judgments but are less closely related to the progress of work.

1.1.1 *Decisions of Buyers*

Decisions of buyers represent the voice of the marketplace: the amount that individuals, libraries, and sublibraries are willing to pay for journals, books, and secondary services. Of course, extensive economic statistics and analysis are required to infer what they are "willing to pay" (always more than what they do pay); also, one must correct for "externalities"–

benefits that society receives but the buyer does not. Still, a plausible calibration can be made in some cases from this type of information. For example, an analysis of available statistics for journals[4] has suggested a direct value of the order of five times their total production cost and an indirect value of similar magnitude. Thus if journal publication of a given small fraction of research were eliminated, the loss to society would be about an order of magnitude more than the production costs saved.

1.1.2 *Time Spent by Users*

Though its significance has not been as widely appreciated as that of free-market prices, the time that users of information services spend in the use of them can be equally valuable for purposes of dollar calibration. Each individual user is perpetually making judgments that balance the value he receives from use of an information service against the value of what he might be doing in the same amount of time devoted to one of his other activities. If one can assume that on the average these judgments are sound—one has to make a similar assumption to calibrate value from the prices buyers are willing to pay—then one can estimate the dollar value of the services *to the users* if one knows the value of the users' time and the (nonlinear) dependence of their ultimate productivity on the time spent in productive work. As before, one must take account of the fact that true social value is not simply the sum of the values perceived by all the individual users. We shall discuss this method of estimating value in more detail in Section 2.4; here it will suffice to mention only one basic input and one result. The basic determinant of the scale of value is the sizable amount, 30 percent to 40 percent, of their total working time that physicists spend in communication activities (see Section 2.1); perhaps one third to one half of this time involves the use of services or facilities for which explicit expenditures must be made. The value of this time must, of course, include loading items representing services that an employer must provide when he hires more employees. Analysis based on such time and cost data shows that any measures that would enable U.S. physicists to do the same amount of communicating via budgetable media in λ percent less time would be worth an amount of the order of (120 to 240)λ/100 million dollars per year (Section 2.4).

1.1.3 *Estimates of Time Saved*

A variant of the user-time input, which was actually employed in a recent study by the American Chemical Society,[5] is based on estimates of the amount of time that users have saved by using certain information services.

This method does not take account of other values they may receive from the services, but it at least gives a lower bound. Unfortunately, no such estimates are available for physicists, and the data for chemists are surely not applicable to them, as their use of secondary services follows a very different pattern (Section 2.1).

1.1.4 *Case Studies of Missed Information*

There have been several studies of instances in which research or development workers have discovered, at a late state in work on some project or after its completion, pre-existing information that would have been valuable if it had been known earlier (for example, Bernard *et al.*[6] and Schilling and Bernard[7]; Brockis and Cole[8]). Rough estimates of economic loss from the missing of this information have been made,[8] amounting to approximately 3 percent of the research and development budget involved; again, these data are not specific to physics.

Now that we have a rough scale for assessing the economic value of possible improvements in communication, it is natural to inquire how great the present cost of communication activities—other than participants' time—looks on this scale. The items to be added are, of course, very diverse: journals, books, libraries, meetings, much travel, telephone bills, some of the cost of lecture rooms, and the like. We shall discuss these in Section 2.3. Our conclusion there and in Section 2.4 will be that the costs of all these different kinds of communication activities and services, which give a total of about $60 million annually, are still not large on the scale of the effective value of the time physicists spend using them, or of the value that one can estimate by willingness-to-pay arguments; for some services, the ratio may be much less than unity.

Similar estimates, which we shall not undertake here, could be made regarding the value of improved communication between physics and other sciences and engineering and in undergraduate and precollege education. We shall say a little about these topics in Chapter 8. However, these brief introductory comments should suffice to indicate why it is important to study the details of the communication patterns of physics and to guide the reader's attention to facts and numbers that will be relevant to attempts to improve this pattern.

1.2 CLASSIFICATION OF THE SUBJECT

Communication has many diverse forms and attributes, and in writing this report it has not been at all easy to decide how to divide the subject into bite-size pieces. There are at least three more or less orthogonal ways of classifying the different kinds of communication activities:

1. We may ask, who is communicating with whom? For example, there is communication of research physicists with one another, there is communication with other disciplines, there is communication of teacher with student. In all cases, we may sometimes wish to focus attention on the supplier of the information, sometimes on the receiver. Moreover, we may compare communication in different countries, or in different kinds of institutions.

2. We may ask, what kind of information is communicated? There are isolated developments on the research front, there is consolidated knowledge, there is knowledge about where to find other knowledge.

2a. Closely related to the kind of information—a possible substitute for it in a classification of communication activities—is the function that the information is intended to serve. There is the function of directing attention to specific items (the "where to find other knowledge" just mentioned); there is use for detailed study and assimilation, which can in turn be for any of several purposes, for example, to bring the receiver up to date in regard to a part of the research front, to supply a needed technique or auxiliary result, to provide background for administrative decisions, and the like.

3. We may ask through what medium the communication takes place. There are books, archival journals, reports, meetings, correspondence, personal conversations, and the like.

The most convenient way of organizing the sections to follow is one based on the classifications 2 and 3 together, with further subdivision of each piece according to 1 when this seems appropriate.

Figure XIV.1 illustrates the way in which classifications 2 and 3 will be used. The three columns correspond to the categories of 2, the rows to those of 3. In each column the rows that are of major importance for that column are crosshatched; those of sizable but secondary importance are lightly shaded. Areas left blank are of lesser, though not always negligible, importance. The heavy contours indicate the coverages of Chapters 3 through 7.

Before beginning these discussions of specific segments of the information-communication picture, however, we need a little perspective. What is the relative importance of the different portions of the picture? What is the topology of their interrelationships—that is, what sequences of conversations, lectures, reading of journals, and the like occur in getting an idea or fact from its originator to a particular user? What are the organizations

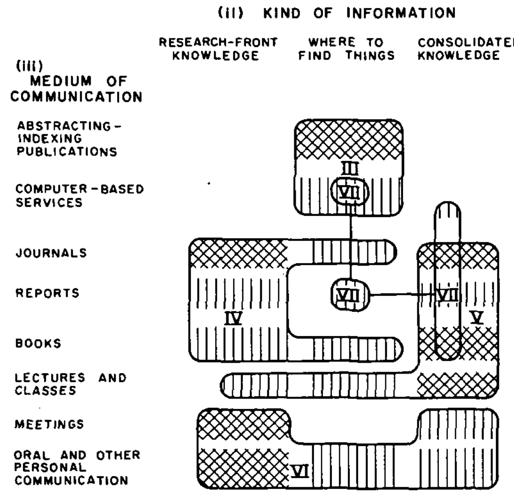


FIGURE XIV.1 Organization of the chapters of this report. In each column, the crosshatching identifies the rows making the major contributions to the particular column; the vertical bars identify the rows making secondary but sizable contributions. Key: III, secondary services; IV, primary publications; V, books, reviews, and compilations; VI, oral and interpersonal communication; and VII, information analysis centers.

that promote communication, and where do the money and other resources come from? These and similar questions will be considered in relation to the overall communication picture in Chapter 2; some of them will, of course, be discussed in greater detail in the later chapters. The very unequal lengths of Chapters 3 through 8 do not reflect vast disparities in actual or prospective importance of the corresponding fields—these are typically of comparable orders of magnitude; rather, they reflect primarily the unequal numbers of studies and amounts of available data.

If we look at acts of communication from the standpoint of a scientist receiving information, an alternative classification scheme, shown in Figure XIV.2, proves useful; we shall employ it frequently in Chapter 2. Here the breakdown uses the same rows (media) as Figure XIV.1, now subdivided a little more finely, but has columns based on the classification 2a (functions), rather than 2, with a further subdivision according to the motivation of the activity. An act of communication, that is, an investment of time by the receiver of information in the use of some one of the media, may be for systematic study and assimilation of a quantity of information (not further subdivided in the figure), or it may be to find out about the existence of items that may later serve as objects of study. In the latter case, there are three possibilities: (a) The receiver's use of the communication medium may have been without a voluntary decision on his part to do so—as when a friend bursts into his office and exclaims "Did you hear . . .?" (b) The receiver may have been on the lookout for interesting things—we call this "browsing"—without having in mind any particular desired item of information. (c) He may have been searching for

MEDIUM		FUNCTION (USE TO WHICH THE MEDIUM IS PUT)			
		DIRECTION OF ATTENTION			STUDY
		GIVER – INITIATED	RECEIVER-INITIATED		
			BROWSING	SEARCHING	
ABSTRACTING-INDEXING PUBLICATIONS	ABSTRACT JOURNALS		○	●	
	CURRENT AWARENESS JOURNALS AND PREABSTRACTS		●		
	CITATION INDEXES			●	
	COMPUTER SERVICES	○ (SDI)		●	
	JOURNAL INDEXES			●	
	RESEARCH JOURNALS		●	● (Cross refs.)	●
	PREPRINTS AND REPORTS	●		○ (Cross refs.)	●
	BOOKS AND REVIEWS		○	●	●
	LECTURES AND CLASSES		●		●
	TALKS AT MEETINGS		●		○
	WRITTEN PERSONAL	●		●	○
	ORAL PERSONAL	●	●	●	○
	OTHER	?	?	?	?

FIGURE XIV.2 Modes in which each of the communication media listed in the first column are used by a receiver of information. A full circle indicates major use, an open circle minor use, relative to the total use of the medium.

a specific item, for example, an absorption coefficient of a particular substance, or an explanation of why the g factor of conduction electrons in InSb is about −50. The figure shows principal and subsidiary ways in which each of the media is used. Although the absence of a circle in any box does not mean that such use never occurs, it implies that it is less important than when a circle is shown, at least at the present time. (Things may change; for example, someday one may be able to browse very effectively on a computer.) Three of the boxes contain explanatory notes: "SDI" refers to selective dissemination of information, a scheme whereby a scientist is automatically sent a small number of titles, abstracts, or reprints chosen by a central agency to match his presumed interests; "cross refs." refers, of course, to the widespread and very fruitful custom of learning about interesting work from its citation in papers one has consulted.

2 Roles and Interrelationships of Different Parts of the Communication Picture

The mechanisms by which scientific information is communicated have for many years fascinated documentalists, sociologists of science, and a few natural scientists. But because most information-gathering activities are private and individual, most of the studies that have been made have been rather unsatisfying; their orientation has been too much determined by what the communications researcher can easily measure rather than by what is ultimately most significant. Even so, some interesting and significant facts have been accumulated. Most of these have to do with the nature of the "last link" in information transfer, that is, either with the mechanism by which a user of information finally gets it or with the relative amounts of time he spends in different kinds of information-gathering activities. We shall discuss such studies in Section 2.1. A few recent studies have examined in more detail the nonuniformity of communication activities and the topology of their coupling to one another; those will be discussed in Section 2.2. In Section 2.3 we shall say a little about organizations and resources, and in Section 2.4 we shall discuss ways of quantifying the value of information services and improvements in them.

2.1 STUDIES OF INFORMATION-GATHERING HABITS

We have just classified, in Figure XIV.2, the ways in which a scientist can receive information through the various communication media. The simplest types of questions we can ask about the various combinations in Figure XIV.2 are such things as: How do physicists and related scientists distribute their time among these various possible communication activities? What is the yield of each of these activities in terms of useful information? Do physicists prefer some of them over others, and by how much? We shall present such data for physicists as are available on these questions, as well as some partial data for other types of scientists. However, these presentations should be preceded by a few words of caution. There is good evidence that different populations of scientists have quite different information-gathering habits. As we shall see below, physicists differ from chemists, and both physicists and chemists differ from biologists; within physics, the variation from one subfield to another, though doubtless less, may be appreciable and has not been measured. Even within a given subfield of physics, there is surely considerable variation from one institution or work environment to another. Particularly noteworthy is the fact that engineers rely much less than scientists on journals for the information they use, and more on oral sources, catalogues, and the like.[9-12] None of the sets of data to be presented can be considered typical of all physics, though features that occur again and again in independent studies are likely to be of general validity.

The effort to synthesize the results of studies of communication habits runs up against further difficulties: Different studies usually use different categories of media; often they pertain to different columns of Figure XIV.2, or different combinations of columns. In Figures XIV.3, XIV.4, and XIV.5, we have tried to sort out these variations, first, by grouping studies of similar nature together in a single figure, second, by placing a few explanatory words on each study at the head of its column, and third, by brackets and explanatory words relating to the table entries.

The most clearly measurable aspect of the use of the various media of communication is the amount of time devoted to them. Figure XIV.3 shows some results obtained in an unpublished 1969 random-observation study by Herring of a sample of U.S. industrial physicists and, for comparison, results obtained in two considerably more extensive studies of chemists. Many necessary explanations and qualifications of the data are given in the footnotes following the figure; for the hasty reader, it will suffice to note only a few of them: The bars in the second column give estimates, based on the observational data of the other columns, of *total* (except in the "oral personal" row—see below) hours per week devoted to reception of information via each of the channels listed in the first column. In many cases, particularly for reading activities, a sizable proportion of these total hours occurs outside the official "working hours" of a scientist's organization. The data in the other columns sometimes refer to working hours only, sometimes to total time, as indicated in the notes. Most of the data of the last column refer to the sum of time receiving and time giving information. Thus a number of somewhat arbitrary assumptions have had to be used to convert the data for chemists into bars for the first column; these are again explained in the notes. In the "oral personal" row, it is almost impossible to separate the giving from the receiving of information, so in all cases the entry here, unlike the others, represents time in both these activities taken together.

It would be very risky, of course, to accept the numbers in Figure XIV.3 as representative of comprehensive averages for either research physicists or research chemists: The two columns for chemists differ somewhat; the two sets of Case Institute data for journal reading by chemists (see note *e*) are not quite consistent; physicists in different subfields or different institutions might well differ from each other by a sizable fraction of the dif-

POPULATION	Physicists (thick bars) Chemists (thin bars)	*Industrial* basic-research physicists[a]	Industrial research chemists[b]	Industrial (not nec. research) + a few university chemists[c]
METHODOLOGY	Estimates from data to right	Random observations + queries[d]	Question-naire	Random observations
MEDIUM / REFERENCE	10 h/wk	Herring[e]	Hannay[e]	Case[f]
Abstracting-*indexing* publications:	*g*		3.08 tot[h]	
Abstract journals		≤0.05	2.59[i]	
Current awareness journals and pre-abstracts		unknown	0.45	
Citation indexes		negligible	0.04	
Computer services		negligible	0.16	presumably none
Journal indexes		in office + elsewhere[d]		
Research journals	*j*	0.63 + 0.40	5.8	1.96[k]
Preprints and reports	*j*	0.79 + 0.31		3.80[l]
Books and reviews	*m*	0.59 + 0.38	2.7	
Lectures and classes	*n*	1.4	unknown	3.68[o]
Talks at meetings	?	unknown	unknown	unknown
Written personal	*p*	≤1.32[q]	unknown	
Oral personal	*r*	≤8.8[s]	unknown	4.08[t]
Other	?	probably small	unknown	small

FIGURE XIV.3 Average hours per week devoted by certain samples of physicists and chemists to the use of various communication media. The bar-length estimates refer to time receiving (or seeking) information only; the precise meanings of the numerical entries appear in the footnotes.

[a] Several score PhD physicists, both theoretical and experimental, in several basic-research departments at the Bell Telephone Laboratories. Mainly but not entirely in solid-state fields.

[b] About 1400 staff members of Corporation Associates of the American Chemical Society selected by company representatives as typical of their "research" chemists. Slightly over half had doctorates; about two thirds were engaged mainly in research; the remainder were mostly occupied with development, a few with administration. Probably at least half were in organic chemistry fields.

[c] A sample of 1500 chemists from 45 industrial and 5 university organizations, systematically chosen from the 172 major metropolitan areas of the United States with the aid of the National Science Foundation and the American Chemical Society.

[d] Time at various activities in an office or laboratory recorded by random observations. Time

ference shown between physicists and chemists. But despite all the fluctuations and uncertainties, there are several conclusions that seem justified; these are to some extent buttressed by the additional data presented in Figures XIV.4 and XIV.5.

1. Physicists, like chemists, spend a sizable proportion of their time, about 15 hours per week, in the reception of scientific information, or in give-and-take oral discussion of it.
2. Physicists spend very little time using abstract journals—much less than chemists. (But we shall see in Figures XIV.4 and XIV.5 that abstracts and title listings are nevertheless a moderately important medium for current awareness.)
3. Reading of journals, preprints and reports, and books and reviews

---

reading in library and at home estimated from combination of these data with estimates by a subsample of subjects that they do about 61% at their offices, 24% at home, 15% in the library. Time in seminars, at meetings, etc., was not recorded.

[e] Herring (unpublished data, 1969); Hannay *et al.*[5]

[f] Most of the data of the Case study (Operations Research Group, Case Institute of Technology) are reported in Halbert and Ackoff;[13] see also Case Institute,[14] Martin,[15] and Martin and Ackoff.[16]

[g] Total for physicists assumed to be a little greater than twice "abstract journals" entry (see Figure XIV.5). Total for chemists assumed to be 0.9 X entry in Hannay column.

[h] Total of the three entries below, which in turn were obtained by multiplying hours of use per week by persons having access to each type of service, as given by Hannay *et al.*,[5] by the fraction of respondents who had such access. Use of these media was roughly evenly divided between use for current awareness and use for search.

[i] Including patent services (0.84 h/week).

[j] Total for chemists assumed to be 0.9 X half the Hannay column for the two rows "research journals" and "preprints and reports" together. This is roughly consistent with an independent Case Institute study.[15] Total for physicists taken as average of Herring column with the 2.4 h/week found for research physicists in an independent study.[14,15]

[k] Apparently refers to both giving and receiving of published information during working hours (assumed to be 40 h/week). The figure 1.96 looks a little low in comparison with the figure 2.0 (2.7) hours for reading of scientific journals by all chemists (or all research chemists) in and out of working hours as measured in an independent study by the same group.[14-16]

[l] Refers to both giving and receiving of unpublished written information during working hours.

[m] Total for chemists assumed to be 0.9 X entry in Hannay column.

[n] Total for chemists assumed equal to fraction of Case column identified by Halbert and Ackoff (1959) as receiving, rather than giving, of information.

[o] This is really the figure identified by Halbert and Ackoff[13] as "oral, non-discussion," which includes giving as well as receiving of information during working hours.

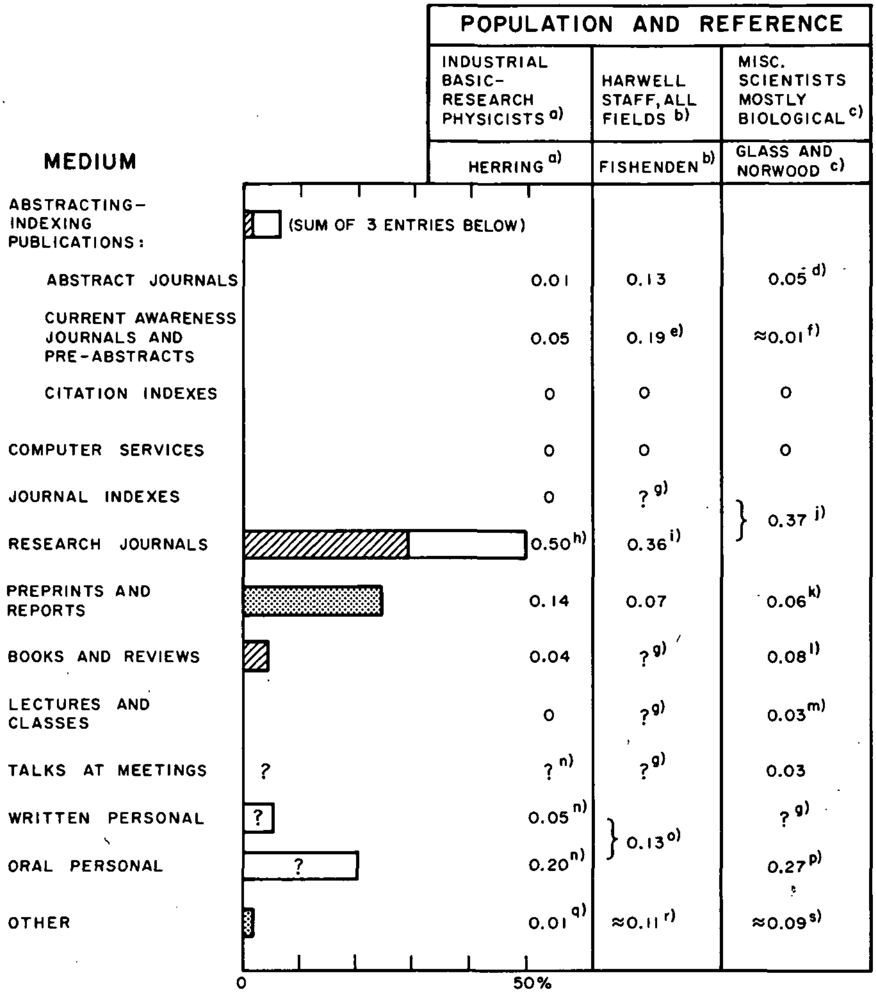
[p] On the basis of fragmentary information[17,18] we have assigned about 0.3 hr/week to both chemists and physicists for reading of scientific correspondence.

[q] Represents reading occasions, in office or laboratory, not assignable to any of the other rows.

[r] Figure for physicists taken as 0.9 X entry in Herring column. Figure for chemists taken to be that in the Case column.

[s] From the frequency of observations of telephone conversations, or of face-to-face conversations with another scientist, in a given subject's office or laboratory, the latter multiplied by two on the assumption that the given subject will talk an equal amount of time on someone else's premises. The factor two is a little too low, since sometimes more than two scientists are involved. On the other hand, some of the conversations were undoubtedly on nonscientific subjects. One should *also add conversations in halls, lounges, etc.*

[t] This is the figure identified by Halbert and Ackoff[13] as "general discussion."



MEDIUM	POPULATION AND REFERENCE: INDUSTRIAL BASIC-RESEARCH PHYSICISTS a) — HERRING a)	HARWELL STAFF, ALL FIELDS b) — FISHENDEN b)	MISC. SCIENTISTS MOSTLY BIOLOGICAL c) — GLASS AND NORWOOD c)
ABSTRACTING-INDEXING PUBLICATIONS:	(SUM OF 3 ENTRIES BELOW)		
ABSTRACT JOURNALS	0.01	0.13	0.05 d)
CURRENT AWARENESS JOURNALS AND PRE-ABSTRACTS	0.05	0.19 e)	≈0.01 f)
CITATION INDEXES	0	0	0
COMPUTER SERVICES	0	0	0
JOURNAL INDEXES	0	? g)	} 0.37 j)
RESEARCH JOURNALS	0.50 h)	0.36 i)	
PREPRINTS AND REPORTS	0.14	0.07	0.06 k)
BOOKS AND REVIEWS	0.04	? g)	0.08 l)
LECTURES AND CLASSES	0	? g)	0.03 m)
TALKS AT MEETINGS	? / ? n)	? g)	0.03
WRITTEN PERSONAL	? / 0.05 n)	} 0.13 o)	? g)
ORAL PERSONAL	? / 0.20 n)		0.27 p)
OTHER	0.01 q)	≈0.11 r)	≈0.09 s)

0 50%

FIGURE XIV.4 Distribution of sources of awareness of useful information over the various media of communication. The Herring and Glass-Norwood studies tabulated sources of awareness of publications cited in papers written by the subjects studied; the Fishenden study[19] tabulated sources of awareness of "useful units of information" (again usually papers or reports). The numerical entries are fractions of all cases (totaling 1.0); the bars in the Herring column present these numbers visually and show their breakdown according to the first three columns of Figure XIV.2; namely, giver-initiated (dotted), browsing (open), and searching (diagonally shaded). Letters refer to explanatory footnotes.

[a] Herring (unpublished data, 1969). About three dozen PhD physicists, both theoretical and experimental, in several basic-research departments at the Bell Telephone Laboratories. Mainly but not entirely in solid-state fields. Interviewed on recollections of how they learned about items

averages typically seven or eight hours per week for research physicists, the three media receiving comparable fractions of reading time. Probably chemists spend a little more time on these. A sizable minor fraction of this reading takes place during other than standard "working hours."

4. At least in large institutions, where opportunities for contact with colleagues are good, physicists are apt to spend about as much time in person-to-person scientific conversations as in all types of reading together.

An additional characteristic of the distribution of time, not shown in Figure XIV.3, concerns the distribution over the columns of Figure XIV.1 or Figure XIV.2. Unfortunately, all that is available is a comparison of chemists' use of secondary services for current awareness with their use for searching, obtained in the Hannay study. This comparison showed that the time devoted to each of the types of secondary services was fairly evenly divided between the two functions.

An alternative to studying the distribution of scientists' time is to study the distribution of sources of useful information. Thus, one can identify an item of information used by a scientist in his work and ask him how he

---

randomly selected by the interviewer from citations in their papers. Items authored by colleagues at Bell were excluded.

[b] Fishenden.[19] Harwell staff members in pure research, applied research, and engineering, 63 in all. Self-identification of "units of information" on diary cards kept for two months. In the present tabulation, items "previously used" have been excluded.

[c] Glass and Norwood.[20] Interviews with 50 scientists on sources of awareness of cited items chosen by the scientists as being of especial importance to them.

[d] Events classified as "through an abstracting service" or "through an indexing service."

[e] Categories "information bulletin" and "library report list."

[f] Events classified as "from a book list."

[g] No relevant category in the results presented, though a perceptive percentage might have been expected.

[h] The diagonally shaded part consists of all items found through reference in another paper, hence perhaps should not all belong to the "searching" category.

[i] A little over two thirds of this entry was listed under "regular reading of current literature," the remainder under "reference in another paper read."

[j] Over four fifths of this entry was listed under "from a journal regularly scanned" or "from a journal subscribed to," the remainder under "from a cross citation in another paper."

[k] Includes reprints as well as preprints.

[l] Categories "from a reference work or textbook" and "from a review article (old work)."

[m] Categories "from a bibliography or material supplied in a course" and "in a formal discussion group."

[n] Category "conversation or correspondence with author" was arbitrarily assigned 2/3 to written, 1/3 to oral communication, but some may well have been talks at meetings. See also Figure XIV.7.

[o] Category "personal recommendation."

[p] Categories "from a co-worker in the same laboratory or department" and "casual conversation," the latter being by far the larger.

[q] A paper refereed.

[r] About two thirds of these events involved various library services.

[s] Entries such as "can't remember," "common knowledge," "by chance."

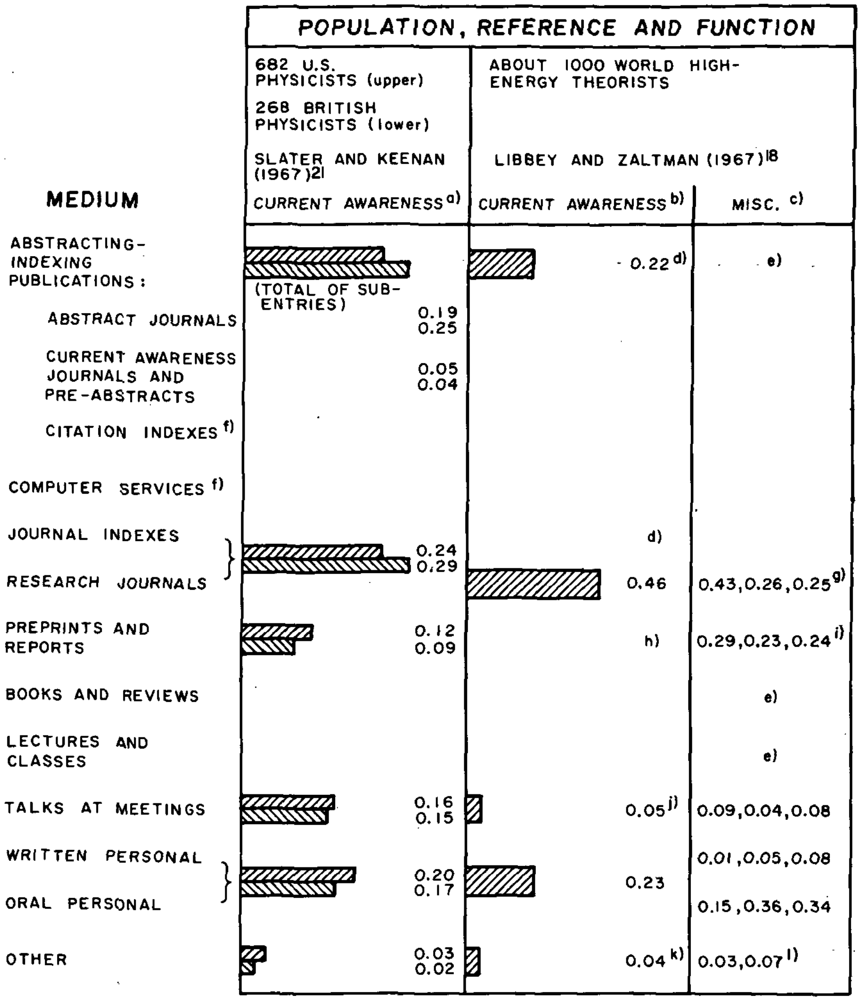


FIGURE XIV.5 Figures of merit for various communication media, as determined from preferences expressed by physicists on questionnaires. Numerical entries are normalized to total 1.00 vertically.

[a] Figure of merit has been defined for each medium by assigning weight 3 to the number of respondents listing this medium as their most important source of current awareness, 2 to the number listing it as second most important, and 1 to the number listing it as also useful. In this way the Slater-Keenan data (their Section 4) for British and American populations could be compared. Slater and Keenan used a more detailed (eight categories) breakdown of preferences to construct figures of merit for their U.S. population; these numbers nearly coincide with those given here. All numbers are normalized to make the sum over media equal one.

[b] Figure of merit is the percentage of respondents selecting each medium as "the most important way of learning of the existence of relevant information" (Libbey and Zaltman,[18] Table 20).

learned it or learned about it. Figure XIV.4 shows the frequency distribution of sources of awareness of such items, as found in a small-scale study of industrial basic-research physicists, and related frequency distributions found in two other studies of scientists in various fields, including physics. The data of Figure XIV.4 give further support to conclusions 2 and 4 above and support these additional noteworthy conclusions:

5. Despite the greater amount of time spent in oral communication, more leads to published items come from the use of journals than from oral sources.
6. Browsing in journals is a major source of useful information, accounting for almost as many leads as the other highly productive category, cross references.
7. Books and reviews provide a surprisingly small fraction of the physicists' leads; their use is more productive for the biologists. A recent study of chemists,[22] not incorporated in Figure XIV.4, also showed them to rely more on books than the physicists depicted here, though still much less than on the primary and interpersonal sources.

Despite some of the differences already noted between the information-gathering habits of physicists and those of other scientists, there are strong similarities in regard to all the above points except the second one. These similarities help to allay fears that data from broader samples of physicists might give a different picture. Fortunately, moreover, there exist data of still another sort with which Figures XIV.3 and XIV.4 can be compared,

---

[c] First figure for medium j is $(\frac{1}{4})\Sigma_\alpha^{(1)} p_{j\alpha}$, where $p_{j\alpha}$ is the percentage of respondents listing medium j as best performing function α, and in $\Sigma^{(1)}$ the function α runs over four values: "general awareness of current state of theoretical high energy physics," "find out who is working in what area or on what problems," "getting up to date in a new area," and "browsing/stimulation." (These four all had fairly similar distributions.) The second figure given is $(1/3)\Sigma_\alpha^{(2)} p_{j\alpha}$, where α runs over "source of specific ideas for work in progress," "source of specific ideas for new work," and "other functions of importance to your work." The third figure is $(1/21)\Sigma_\lambda^{(3)} f_{j\lambda}\,(7-\lambda)$ where $f_{j\lambda}$ is proportional to the fraction of respondents assigning rating λ to medium j ($\lambda = 1$ means very important, $\lambda = 7$ nearly useless, etc.).

[d] Journal indexes were grouped with abstracting-indexing publications in this study.

[e] These communication media were not included among the alternatives to be chosen; some contribution to these entries might therefore be expected from the "other" total in the last row.

[f] Availability at the time of these studies was surely too limited for use to be appreciable.

[g] Including reprints.

[h] Not included among the choices offered; there should obviously be a significant entry in this position, which might be drawn from the figures given for research journals (identified in this study as "regular scanning of literature"), oral and written personal (identified as "colleagues") and other.

[i] "Manuscripts," "technical reports," and "copies of oral presentations."

[j] "Programs and proceedings of scientific meetings." It is not clear whether some of the entry for oral and written personal (identified as "colleagues") should be transferred to here.

[k] May include some use of preprints, etc.—see note *h*.

[l] May include contributions for the rows labeled with note *e*.

and that are, in fact, available for some quite different samples of physicists. These are data on subjective preferences among the different communication media, or subjective evaluations of their relative importance. These are of some value despite the ambiguity of quantitative measures of them and possible systematic tendencies to erroneous judgments. Figure XIV.5 shows the results of two such preference studies. Although not shown in the figure, another study by Martyn[23] leads to similar conclusions; it also compares scientists in different disciplines.

The first study, by Slater and Keenan,[21] is based on a questionnaire concerning current-awareness habits sent to a large sample of British and American physicists in 1966. Although the physicists approached were selected by random sampling of (a) the membership of the American Institute of Physics, The Physical Society, and the Directory of British Scientists and (b) the physicists in the National Register of Scientific and Technical Personnel, the percentage of usable returns (30 percent and 26 percent, respectively) was small enough to cast doubt on the representativeness of the data. As the results have to do mainly with current awareness, they pertain only to the "browsing" and "giver-initiated" columns of Figure XIV.2. No questions were asked about time or frequency of use of the various channels, but the respondents were asked to rate the channels in estimated order of importance. From these rankings we can construct a somewhat arbitrary but doubtless meaningful figure of merit, as explained in the notes.

The other preference study summarized in Figure XIV.5 is based on answers to a questionnaire mailed to nearly 4000 presumed high-energy theorists throughout the world, with usable (for the present purpose) returns from a little under 1000, estimated to represent over one third of the true population of high-energy theorists.[18] Respondents were asked to identify which of several communication media best fulfilled each of a number of functions for them, functions often compounded of current awareness, search, and study. The current-awareness function and three composites of the many others have been chosen for tabulation in Figure XIV.5.

Despite the arbitrariness of the entries—most seriously, the fact that grouping several communication media together could well yield a higher figure of merit than the total for those media offered separately—the resemblance of Figure XIV.5 to Figures XIV.3 and XIV.4 is strong. The role of abstract journals seems more important in Figure XIV.5 than in the earlier figures, especially in the Slater-Keenan column; this is doubtless due in part to the fact that only current-awareness activities are being looked at and may be enhanced by a tendency for those who returned the Slater-Keenan questionnaire to be people who use abstracting and current-awareness services frequently. The patterns of Figure XIV.5 are reasonably consistent with statements 3 through 6 above. They provide particular support for 6, the importance of browsing in journals, which in the Slater-Keenan study got top preference from all of the eight groups of physicists studied (U.S. and U.K. physicists in four types of institutions). In addition, they suggest the following conclusion, not obvious from the earlier data:

8. Talks at meetings provide an important source of current awareness, probably comparable in average importance to preprints, and somewhat less important than journal browsing and personal conversations. They are less useful as a source of detailed specific information.

A final point is very important. Many studies of all types of scientists and engineers[24–26] have shown:

9. Easy accessibility is a major determinant of the extent to which any formal or informal information resource is used. Physical proximity to the user's office or laboratory is a very important component of this; an additional hundred feet of distance can retard use quite noticeably.

Of the various studies of use of information channels by scientists and engineers, we have touched on only the few that pertain specifically to physics, or to sciences similar to physics. References to recent studies in other disciplines can be found in *Annual Review of Information Science and Technology* (C. Cuadra, Editor, Encyclopaedia Britannica, Inc., Chicago, annual); see particularly the reviews by H. Menzel (Vol. 1, 1966), S. and M. Herner (Vol. 2, 1967), W. J. Paisley (Vol. 3, 1968), and T. J. Allen (Vol. 4, 1969). Earlier studies of some interest include those of Kotani[27] on Japanese scientists, Herner[17] on American medical scientists, Project on Scientific Information Exchange in Psychology[28,29] on psychologists, Martyn[23] on British scientists, and Columbia University[22] on polymer chemists. A summary of many studies has been given by Vickery.[30] Hagstrom[31] has published some interesting statistics on publication by academic scientists. Additional information on physicists, but pertaining to only one or a few of the communication media, will be discussed in subsequent appropriate sections of this chapter.

2.2 INDIVIDUAL DIFFERENCES AND THE TOPOLOGY OF COMMUNICATION

The picture we have developed so far is incomplete: It has dealt only with averages over many scientists in a given area, and only with one of the functions (columns of Figure XIV.2) at a time, most often a direction-of-attention function. It would be helpful to know how great are the differences between the communication patterns of different individuals in a given field and what sequences of communication acts are involved in getting useful information from its originator to an ultimate user. Unfortunately, very little has been done on the study of these questions in the physics community, so our remarks will have to be based in part on studies of chemists and engineers and in part on very sketchy observations of physicists.

Purely individual variations seem to have been very little studied, even in other fields. In the Case study,[13] about 1000 chemists were each observed at approximately 18 random moments. For each of four types of scientific communication (general discussion, oral nondiscussion, written unpublished, and written published), and also for the sum of these four activities, the report gives the mean percentage of time devoted to such communication and also the maximum and minimum percentages observed among the 1000 subjects. It is easy to make a rough check, using the binomial distribution, of the reasonableness of the hypothesis that the population was uniform in its communication habits and that the variations of the maxima and minima from the mean were due purely to the random occurrence of the events in question among the 18 observations. When this test is applied to each of the four types of communication individually, the minimum times (zero in every case) turn out to be consistent with the hypothesis; the maximum times are also roughly consistent except in the case of published material, for which the observed maximum (18.4 percent) was rather *less* than the expected maximum for a mean of 4.9 percent. For the total time on all four types of communication together, the minimum (15 percent) was a little *more* than would be predicted from the mean (33.4 percent), while the maximum (64 percent) was roughly consistent. Departures from the mean less pronounced than expected by chance suggest that the 18 or so observations on a given individual are not quite statistically independent; this fact weakens slightly the main and surprising conclusion from these data, namely, that variations from individual to individual may not have been very wide in this population.

Despite this rather negative, or at least inconclusive, result, the Case Institute studies did reveal real variations among types of individuals and types of institutions. One expects communication habits to be affected by different institutional surroundings (availability of colleagues in the same field or subfield, geographic convenience of the library, and the like) and indeed, Martin and Ackoff[16] did mention a difference of more than a factor three in the time devoted to literature reading between chemists in two extreme groups of eight and six companies, respectively. They also noted that the relative amount of time spent with written communication as opposed to oral was greater for industrial than for academic chemists. But even within a given institution, there are not only variations of a purely individualistic nature but also systematic variations with work activity and talent. Thus, Martin[15] found that among both physicists and chemists the scientists active in research spent over twice as much time reading journals as nonresearch scientists. Other studies by von Zelst and Kerr[32] and by Maizell[33] have shown clear positive correlations between amount of journal reading by different individuals and various measures of their productivity.

A recent study,[34] which unfortunately covered both nonprofessional and professional personnel, gave the results shown in Figure XIV.6 for

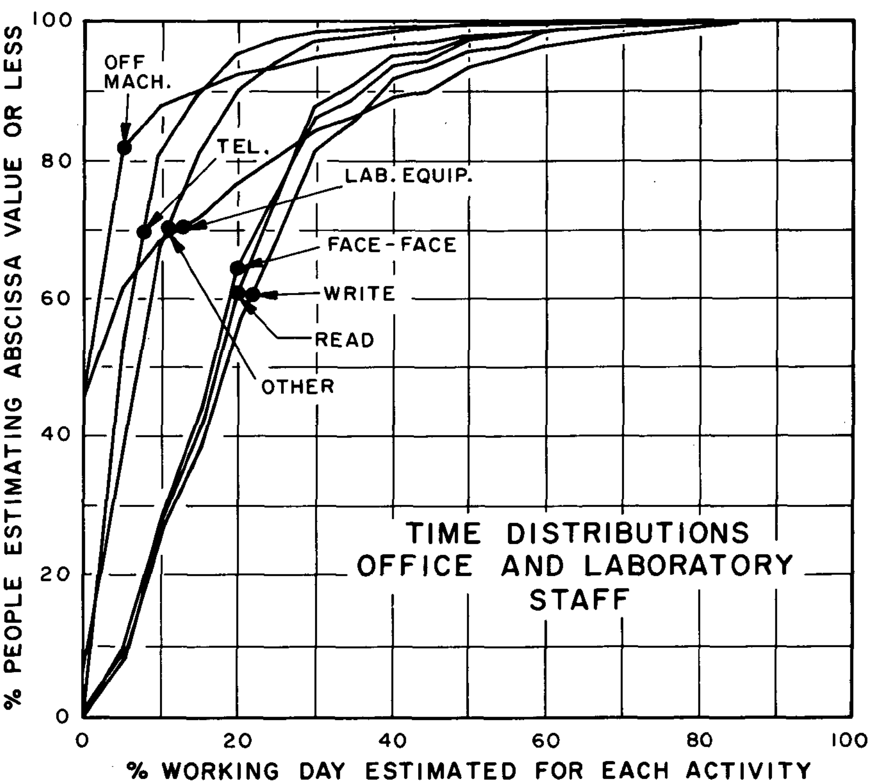


FIGURE XIV.6 Spread of time distributions for various activities, as reported by all office and laboratory workers (professional and nonprofessional) of a large industrial research laboratory.[34] Dots indicate mean values.

the individual variations in time devoted to various types of activities. Although these results need adjustment for the fact that probably over half of the subjects were nonprofessional and for the fact, uncovered in another part of the study (see also Hinrichs[35]), that the responses systematically overestimated reading and underestimated talking, they can still be used to give upper bounds for the range of individual variations among scientists and engineers.

Some recent studies on communication patterns in industrial laboratories not only have revealed differences between individuals but have shown how some of these differences play a very important role in the functioning of a scientific or technical group.[26,36,37] These studies have shown that in typical organizations there exist special individuals—called "gatekeepers"—who serve as focal points for information originating outside the organizations. Specifically, a person who is chosen by an unusually large number of his colleagues (within the organization) for frequent participation in technical discussions is characterized as a gatekeeper. The studies have shown that the gatekeepers so identified differ considerably from the remainder of their colleagues in that, on the average, they read many more scientific and professional periodicals and have many more information-producing contacts with friends outside their organizations.

Studies such as the ones just described merely whet one's appetite to know how useful information diffuses in the community of research physicists. We have undertaken a small-scale study of one portion of this community, which will serve at least to illustrate the type of conclusions that a more extensive study might be able to yield, though the specific conclusions to be described presently might not be valid for categories of physicists other than the one studied. The group in question consisted of highly productive basic-research physicists, mostly in solid-state fields, at a large industrial laboratory. Two recent papers written or coauthored by each physicist were selected at random, and from each paper several ideas and facts were selected, again at random, that (a) contributed perceptibly to the arguments of the paper, (b) did not seem to have originated with the author or authors of the paper, and (c) did not seem to be such common knowledge that they could be expected to be known by the average PhD in the field. The physicist was then interviewed and asked:

1. How the specific information he actually used came to him, that is, through reading (what?) or listening (to whom? where?).

2. How he was cued to get it. (For example, did someone refer him to a paper? Send him a preprint? Had he undertaken a search for this type of information?)

3. Did he know how the information, or knowledge of its existence, got to be the source or cue cited in 1 or 2?

4. How much time elapsed between the original discovery of the fact or idea and its apprehension by the interviewee?

Figure XIV.7 displays the distribution of communication chains found in this study. Some obvious conclusions may be drawn:

1. Direct transfer of the information from the originator to the user or (more rarely) transfer by an initial preprint mailing occurred in about a fifth of the cases. Even less common was transfer in which two or more intermediate repositories took part (for example, a journal article and a

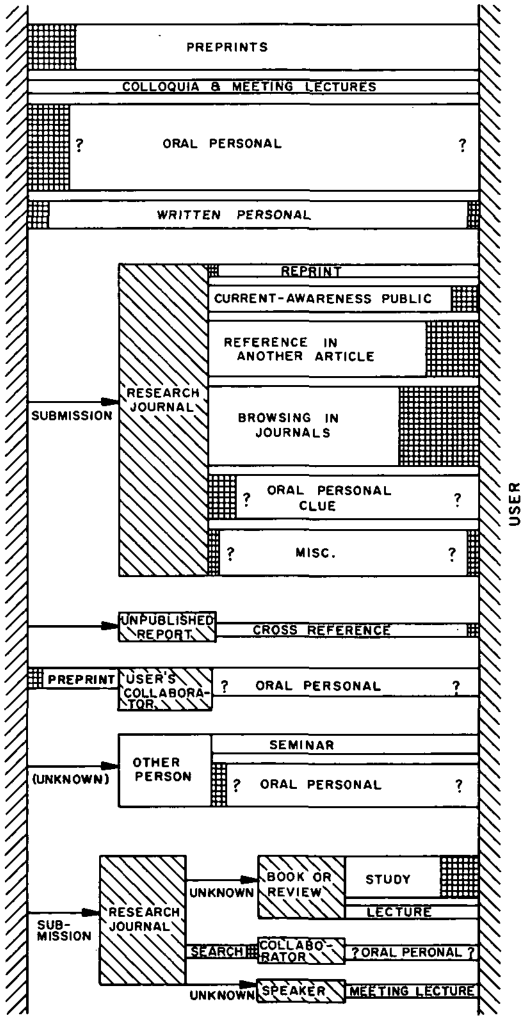


FIGURE XIV.7 Distribution of transmission chains for important items of information used by a sample of condensed-matter physicists. Diagonal shading identifies "repositories" in which information resided for an appreciable time between transmissions; unshaded rectangles identify different types of clues that led to assimilation of information from use of these repositories. Vertical widths of the repositories are proportional to number of cases of each type in the sample; arrows are used in place of rectangles when classification by clues is inappropriate. Rectangular crosshatching at the right or left ends of the rectangles designates the numbers of cases in which the initiative was with the giver (left) or the receiver (right) of the information; question marks indicate further cases in which the source of initiative was in doubt. In some cases (for example, lectures), the question of initiative is meaningless and the crosshatching is omitted.

book). Thus the great majority of the cases involved a single intermediate repository between the originator and the user, most often a journal article but sometimes just the mind of a third person.

2. The actual acquisition, by the user, of the specific information of importance to him occurred through reading a primary journal article in nearly half the cases. In a sizable minority of the remaining cases, a journal article constituted one of the steps in the transfer of information from originator to user.

3. In nearly one third of the cases, the specific information of importance was acquired by the user in oral conversation with a colleague, most often the originator himself, and in other cases frequently a collaborator of the user. In a sizable minority of the remaining cases, an oral-personal link played a part, most frequently by providing a lead to printed material.

4. Over half of these oral-personal links were giver-originated; giver-initiated written-personal, preprint, and reprint links also occurred. In total, giver-initiated links occurred in over one third of the transfer chains. There were no conspicuous differences between the communication patterns described by experimentalists and theoreticians or between those associated with experimental and theoretical information.

An interesting datum collected in the interviews was the time interval between the first publication or communication of the item of information by its originator and the receipt of the information by the interviewee. Figure XIV.8 shows the distribution of these time intervals; it can be summarized in the following statements:

5. In about half of the cases, the time interval was no more than about a year, but in most of the remaining cases it was more than five years.

6. For the cases with a long time interval, acquisition of the information via primary research journals was predominant; in a sizable minority of the remaining cases, the information was rediscovered, and contact with the original discovery was made only after search or inquiry. Books and reviews played a surprisingly small role in these long-interval cases.

2.3 ORGANIZATIONS AND RESOURCES FOR THE COMMUNICATION OF INFORMATION

A glance at the communication media listed in the first column in each of Figures XIV.2–XIV.5 shows that some involve vast programs of national or international organizations, some involve commercial, industrial, or non-profit activities of modest scale, and some are at the level of informal

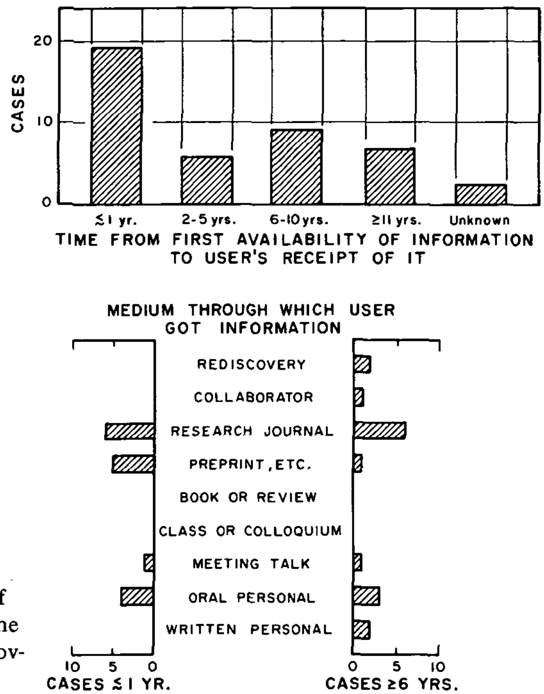


FIGURE XIV.8 Distribution of time intervals, when known, in the use of information in the cases covered by Figure XIV.7.

everyday activities of individuals. The largest operations are easy to enumerate, but as they are sometimes interdisciplinary and are usually circulated throughout the world, it is not always easy to decide how much of them should be assigned to physics, or to the United States. For the activities of commercial organizations, which may be of large or medium scope, data on costs, circulations, and the like are not easily obtainable. For such small-scale activities as seminar talks and personal correspondence, not only are data unavailable, but it is often hard to draw a line between communication and other activities. Thus our estimates of U.S. resources invested in communications will have to be very crude.

Still, a rough perspective is better than none. So we shall try to enumerate the leading organizations contributing to the various communication media we have been discussing, with a rough measure, monetary if possible, of the size or importance of each. In many cases the data are taken from later sections of this chapter, where they are discussed more fully. Let us start with the organizations having the largest activities relevant to communication in physics in the United States:

1. The American Institute of Physics (AIP), which publishes over 80 percent of the primary physics journal literature in the United States,

produces translations of the leading Russian journals, publishes an appreciable part of the review literature, and provides a number of secondary services. These information activities consume the major part of AIP's operating budget (67 percent of $10.0 million in 1969). The funds involved come primarily, and in comparable amounts, from the sale of subscriptions and from page charges; extensive developmental activities in the information field have been financed by NSF grants. The organization of AIP is centralized under a Director, who, together with the Treasurer and Secretary, is responsible to a Governing Board consisting of members named by the seven member societies, plus a few elected at large by their combined membership. Further contact with the physics community is maintained by a considerable number of advisory committees, including one on publications and one on the information program.

2. The Institution of Electrical Engineers (IEE), in London, which produces *Physics Abstracts* and the title listing *Current Papers in Physics* as well as two abstract journals in electrical engineering areas. As about 45 percent of the subscriptions to *Physics Abstracts* are from the United States, its production and distribution represents about $0.3 million annual expenditure for this country. The governance of *Physics Abstracts* is centralized under a Director responsible to the central administration of IEE, which in turn is responsible to a Board and through them to the Institution's membership. An Advisory Committee has representation from the American Institute of Physics and The Physical Society, as well as from IEE.

3. The U.S. Atomic Energy Commission, which produces *Nuclear Science Abstracts,* supports a variety of specialized information centers, subsidizes some conferences and book production, and is taking a number of steps to promote international cooperation in the handling of nuclear information. The investment of resources in *Nuclear Science Abstracts* is similar to that in *Physics Abstracts* (item 2 above), but the bulk of the cost is met centrally, rather than through subscriptions.

4. Chemical Abstracts Service. *Chemical Abstracts,* though ostensibly devoted to chemistry, has such excellent coverage of a large part of the physics literature that it is extensively used by physicists. Physicists probably make very little use, however, of the other information services provided by this organization (title listings, computer tape searches, special secondary services for particular subfields, and the like). Thus, although the total budget of Chemical Abstracts Service is very large ($18 million in 1971), service to U.S. physics should be reckoned as only a small fraction of the part (about $13 million) allocable to *Chemical Abstracts* itself. Chemical Abstracts Service is administratively responsible to the Executive Director of The American Chemical Society, and through him to the elected Board of Directors of the Society.

5. The National Bureau of Standards, which operates the National Standard Reference Data System (NSRDS) with its extensive publication of critical data collections and which supports various specialized information centers. The budget of NSRDS is currently about $2.4 million a year; perhaps half of its work could be considered allocable to physics.

6. The Institute for Scientific Information, a commercial organization in Philadelphia, which produces the *Science Citation Index,* the title listing *Current Contents,* and other indexes of less interest to physicists, and which offers a flexible computer-based selective alerting and dissemination service to individual subscribers. Data on the number of subscriptions to these services are not publicly available, though ISI has estimated that *Current Contents, Physical and Chemical Sciences* is scanned by about 40,000 physicists, chemists, earth scientists, mathematicians, and others throughout the world. It seems safe to say that the share of the resources entering into these services that is allocable to U.S. physics is of the same order as *Chemical Abstracts.*

Besides these individual organizations with large programs and budgets, the following classes of organizations merit special mention:

7. The scientific societies devoted to physics or to various branches of it. Some of these societies, most notably the largest one, the American Physical Society, provide editorial management and policy direction to journals published for them by the American Institute of Physics (which, of course, also publishes many journals not responsible to a particular society). Another major function that all the societies fulfill is the conduct of scientific meetings. As the finances of the publications are covered largely under 1 above and meetings are financed largely by the participants, there is little in the way of dollars to report here, though probably a few hundred thousand dollars of society dues might properly be assigned to these information activities.

8. The commercial publishers. These produce nearly all the book literature of physics. Excluding popularizations and those textbooks that are used almost entirely by students, and with some rather arbitrary estimates of the boundaries of "physics" and of sales volumes (see Section 5.5), we can estimate the annual expenditure of U.S. physics on books devoted to research in physics and its applications at roughly $11 million. Although U.S. publishing firms publish only a small fraction of the journal literature of physics (see Chapter 4), European commercial publishers account for a sizable fraction of it. We estimate the expenditure of U.S. individuals and institutions for commercially published journals to be of the order of $8 million.

9. Universities and other research institutions. These support external

communication in two major ways, namely, their support of attendance by their staffs at meetings and other such occasions and their maintenance of libraries. As is shown in Chapter 6, the former probably involves a national expenditure of the order of \$8 or \$10 million. As for the expense of library maintenance, part of this represents acquisition costs for journals, books, and secondary services and is included in the figures we have already given on production costs in items 1, 2, 3, 4, 6, and 8; a part, however, represents costs of storage, cataloguing, and maintenance of general library services. These costs are probably a little more than the acquisition costs (see, for example, SATCOM,[4] Appendix Section III D.4). If we allot approximately half of the book and journal production costs to library purchases, and most of the cost of abstract journals, we can estimate the amount of additional library costs allocable to physics research in the United States at about \$20 million per year.

Several other aspects of the support of communication by these institutions deserve mention but become increasingly nebulous. One is the hidden subsidy of primary journals through refereeing, support of editors and their offices, and like services; according to earlier rough estimates (SATCOM,[4] pp. 76–77), this hidden subsidy has a value of the order of one tenth the production cost of the journals, that is, of the order of \$0.7 million per year for U.S. physics. Another aspect is internal communication: conference rooms, internal memoranda, and the like. Still another is the typing of letters and manuscripts and the preparation and mailing of preprints.

Figure XIV.9 gives a summary of the estimates (sometimes extremely crude) developed in the above paragraphs for resources, allocable to U.S. physics research, expended on various aspects of communication. Note that we have intentionally omitted including in these the value of the time physicists devote to communication activities; this value has been estimated in Section 2.1 and Figure XIV.3. In the following subsection, we shall discuss the inferences that can be drawn from a comparison of the value of this time with the cost of providing the information services used.

2.4 POTENTIAL VALUE OF IMPROVEMENTS IN INFORMATION AND COMMUNICATION SERVICES

We are now in a position to carry through the calculation referred to in Section 1.1. We want to estimate the economic value of a hypothetical change in the efficiency of the network of information and communication services. To make this estimate realistically, we must take account of several factors: (a) The time spent in communication activities subtracts

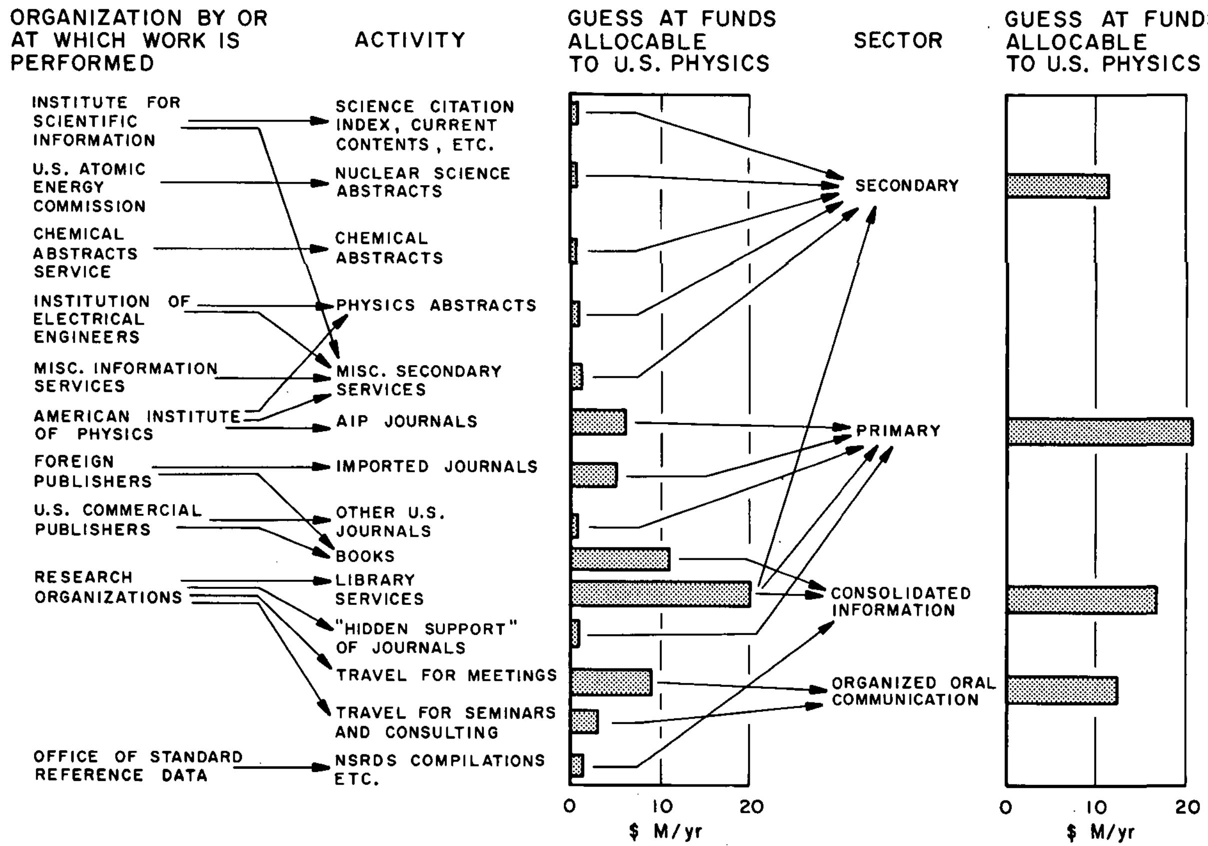


FIGURE XIV.9 Organizations and activities relevant to identifiable organized support of communication in physics (excluding education), with rough estimates of the dollar value of expenditures on behalf of U.S. physics.

from the time available for other work. (b) The efficiency of productive work increases with increasing time t_i devoted to acquisition of relevant information, at least when t_i is small, but at a rapidly decreasing rate as t_i becomes larger. Each research worker must judge for himself what his t_i should be to optimize his total useful output; we shall assume that on the average these judgments are correct ones for the conditions actually obtaining. (c) The function relating the yield Y of useful work to the difference $(t - t_i)$ between total working time t and time spent in communication is nearly always superlinear; for example, the efficiency of research work usually increases as the time t_r available for it increases, and t_r itself is usually of the form $(t - t_i - \tau)$, where τ is time devoted to non-research-related activities.

Let us therefore write, for the useful yield from a nominal working time t,

$$Y/t = \phi(1 - \bar{t}_i)\, g(\bar{t}_i), \tag{2.1}$$

where $\bar{t}_i \equiv t_i/t$, ϕ is a function of $(1 - \bar{t}_i)$ whose normalization may be arbitrary but which does not depend on the nature of information-seeking activities, and g is an efficiency function, dependent on the latter activities. Our assumption that collective individual choices optimize Y/t under present conditions implies that

$$\phi(1 - \bar{t}_i^{(0)})\, g_0'(\bar{t}_i^{(0)}) = \phi'(1 - \bar{t}_i^{(0)})\, g_0(\bar{t}_i^{(0)}), \tag{2.2}$$

where g_0 is the present form of g. Now suppose that some change is instituted that replaces g_0 with g_1. Although the new optimum value of $\bar{t}_i$ will differ from $\bar{t}_i^{(0)}$, this difference will not affect Y to the first order, because Y was initially stationary with respect to $\bar{t}_i$, as manifested in Equation (2.2). Thus the change in Y will be given, to the first order in $(g_1 - g_0)$, simply by

$$(Y_1 - Y_0)/t = \phi[1 - \bar{t}_i^{(0)}] \left\{ g_1[\bar{t}_i^{(0)}] - g_0[\bar{t}_i^{(0)}] \right\}. \tag{2.3}$$

As an important special case, let us consider an improvement that reduces the time t_i, necessary to get a given amount of useful information, by a factor $1/(1 + \lambda)$. This means that

$$g_1(\bar{t}_i) = g_0[(1 + \lambda)\bar{t}_i], \tag{2.4}$$

so that $g_1 - g_0 = \lambda \bar{t}_i g_0'(\bar{t}_i)$ to first order. Then Eq. (2.3) gives, with use of Eqs. (2.2) and (2.1),

$$Y_1 - Y_0 = (Y/t)\ (\phi'/\phi)\, \lambda t_i, \tag{2.5}$$

that is, the improvement in useful yield is equivalent to that which would result from adding to the number of man-hours worked an amount exceeding λt_i by the factor ϕ'/ϕ. As ϕ is surely an upward-curving function of its argument $(1 - \bar{t}_i)$, ϕ'/ϕ is appreciably greater than $(1 - \bar{t}_i)^{-1}$. Since we have seen in Section 2.1 and Figure XIV.3 that $\bar{t}_i$ for physicists is probably about 0.3, we have concluded that *a time-saving innovation in scientific communication resources adds as much to productive output as an increase in total working time of rather more than one and one half times the communication time saved, probably two to three times.*

One more factor must be considered, however, to put a dollar vallue on the innovation: namely, at what rate should the value of physicists' time be priced in applying the result just derived? Should one use bare salaries, salaries loaded with various overhead expenses, or something in between? A reasonable answer is provided by dividing expenditures on behalf of basic and applied research in physics into three categories:

$F_i \equiv$ expenditures explicitly identifiable with provision of information or communication services, that is, those discussed in Section 2.3 and Figure XIV.9;

$F_w \equiv$ those other expenditures that are more dependent on the amount of research done than on the number of people involved in doing it;

$F_p \equiv$ those other expenditures that depend more on the number of people working than on their productivity.

Thus F_p would include unloaded salaries, employee benefits, office and most laboratory space, most administrative overhead, part of secretarial and telephone service, and the like; F_w would include such things as shop facilities, equipment and supplies, and purchasing administration. Studies of typical organizational and project budgets suggest that $F_p \approx 1.55\ F_s$, where F_s is the base (unloaded) salary budget. We propose that the equivalent time saving discussed in the previous paragraph should be converted to dollars at the rate F_p/F_s times the mean bare salary.

One rather sizable uncertainty remains: How many scientists should be reckoned as effective users of the physics information and communication services we have been discussing in Section 2.3, whose annual cost we estimated at \$60 million? These services are oriented mainly toward basic and applied research; one extreme assumption would be to count each physicist as a fraction of an "effective user" equal to the fraction of his time that is nominally devoted to research. With the approximate figure of half a billion dollars for the annual U.S. investment in physics research from all sources and a guess at 30 percent of this as representing the bare-salary component, this philosophy gives about \$150 million per year as a measure of the bare salary F_s to be used in the preceding paragraphs. But this estimate is surely too low: Meetings, research journals, advanced books, and the like benefit many who are not engaged in current research. An alternative estimate of the number of effective users, doubtless too high, would be to take this number to be equal to the numbers of "physicists" in the National Register, that is, about 32,000 in 1968, or 36,000 in 1970. With a mean annual salary of about \$16,000 this extreme approach gives an F_s of about \$570 million per year. The appropriate value for the "total salaries of effective users" thus should lie between \$150 million and \$570 million; the studies of Martin[15] and others, which we discussed in Section 2.2, suggest that a value of the order of two thirds of the upper figure might be reasonable. Thus we guess $F_s \approx$ \$300 million to \$400 million. The corresponding partially loaded salary total, discussed in the preceding paragraph, would then be $F_p \approx$ \$460 million to \$620 million.

Let us for the moment lump all kinds of scientific communication together, including informal conversations and other impersonal interaction, and take from Figure XIV.3 and the associated discussion a value about 0.35 for $\bar{t}_i$, the fraction of total time devoted to such communication. Then the effective value of this time is about 0.35 $F_p \approx$ \$160 million to \$220 million per year; saving a small fraction λ of it will be worth an amount larger than this by the last bracketed factor in Eq. (2.5), which we estimated at 2 to 3, that is, an amount in the range \$300–600λ million per year. A saving of a fraction λ of the time spent on a particular group of communication channels (for example, the formal ones) with no other changes thus has a value of the same form, with the numbers (300 to 600) reduced by the fraction of total communication that is spent on the channels in question (~0.4 for the formal channels).

The simple treatment we have given takes inadequate account of the cooperative nature of information transfer—the way the different channels feed each other—shown in Figure XIV.7. Here we shall mention only briefly one important effect that further enhances the value of improvements in information services. Improvement of a formal channel—journals for example—improves the knowledge and awareness of those with whom a given scientist interacts informally, and hence improves the effectiveness of the "grapevine."

3 The Secondary Services

Having thus summarized the overall picture of communication in physics, we next present a more detailed look at each of the major components of the picture. Although the order in which these different components are discussed is to some extent arbitrary, there are good reasons for beginning with abstracting and indexing publications, since these are the most copious source of data about the primary journal literature, which according to Figures XIV.3–XIV.5 (especially Figure XIV.4) is probably the most important of the analyzable components of the communication picture. They also provide considerable, though less complete, information about some of the other components. No other source of data about the literature of physics is as detailed. But to make intelligent use of data from abstracting and indexing publications, we must understand their characteristics and limitations. Hence we shall start with a look at these publications and, for completeness, at other types of secondary services as well.

We shall begin, in Section 3.1, with a survey of the many existing secondary services relevant to physics, so that we can sort out the general from the specialized, those accessible to the bulk of American users from those in unfamiliar languages, and the like. Then, in Section 3.2, we shall survey statistics on the bulk and coverage of the leading abstracting journals and related services and their changes with time. Section 3.3 will discuss their time lags, distribution, and availability to physicists; Section 3.4, their information content, indexing, and organization; Section 3.5, their use; and Section 3.6, some of the technical problems of producing them and their costs.

Many interesting facts about abstract journals and other literature services have been summarized in a review by Anthony *et al.*,[38] to which we shall refer from time to time; a shorter article by Keenan[39] is also worth noting, as is a broader book by Whitford.[40]

3.1 NUMBER AND VARIETY OF SECONDARY SERVICES

In a recent study,[41] some 69 secondary services, serving physics or its subfields or peripheral areas, were identified and discussed. Among these, there are only four abstract journals that undertake comprehensive coverage of published articles from all parts of the world in essentially all the subfields of physics. These are listed in the top half of Table XIV.1. Note that only one of these, *Physics Abstracts,* is in English. There are several other publications that undertake a reasonably comprehensive coverage

TABLE XIV.1 Abstract and Title Journals with Comprehensive Coverage of Journal Literature in All Major Areas of Physics[a]

Publication	Publisher	Language	Entries[b] (thousands)
Abstract journals			
Bulletin Signaletique (5 physics-related sections)	CNRS, Paris	French	85 (1968)
Physics Abstracts	IEE, London	English	50 (1968)
Physikalische Berichte	Verband Deutscher Physikalischer Gesellschaften Braunschweig[c]	German	36 (1968)
Referativnyi Zhurnal Fizika	Academy of Sciences U.S.S.R., Moscow	Russian	79[d] (1968)
Title journals			
Bibliografia Brasileria de Matematica e Fisica	Ins. Brasiliero de Bibliografia e Documentacio, Rio de Janeiro	Original	–
Current Contents, Physical Sciences	Institute for Scientific Information, Philadelphia	English	104 (1968)
Current Papers in Physics	IEE, London	English	34 (1968)
Current Physics Titles[e]	AIP, New York	English	–

[a] Data are from Terry and Cooper[41] except where otherwise referenced.

[b] The significance of these numbers is somewhat arbitrary, as they reflect not only the intensity of the coverage (number of journals scanned and the like) but also the coverage or noncoverage of report literature and like material and, especially, the extent to which nonphysics areas are included. Thus, for example, the coverage of *Referativnyi Zhurnal* probably does not differ greatly from that of *Bulletin Signaletique*, but, whereas in the former all of basic physics is included in the subdivision Fizika, which we could therefore list by itself in the table, several subdivisions of *Bulletin Signaletique* are required to cover all of basic physics, so the entry in the last subdivision is a total for five sections, some of them rather peripheral to physics. Again, *Current Contents, Physical Sciences* covers a number of sciences together and so has a very large entry.

[c] Altered from the Cooper-Terry listing.

[d] Approximate count, 1968.

[e] New publication.

of all physics but that list titles only. These are enumerated in the lower half of Table XIV.1. Two further categories, between which it is hard to make any sharp distinctions, are abstract and title journals that undertake comprehensive coverage of articles in specific subfields of physics and those devoted to nonphysics fields that overlap physics. Publications of these types are listed in Table XIV.2 (abstract journals) and Table XIV.3 (title journals); the listings of journals in peripheral fields are not intended to be complete, as the decision that such fields do or do not overlap physics is rather arbitrary. Note that some of the publications listed in Tables XIV.2 and XIV.3 are abstract or title listings in a special section of a primary journal, rather than journals devoted entirely to the secondary function.

Tables XIV.1, XIV.2, and XIV.3 have listed only about half of the secondary services identified by Cooper and Terry. Of the remainder, some seem to have very incomplete coverage, as judged from a comparison of the number of items listed with the field to which they are devoted; some cover only work published in a particular region (for example, the Soviet Union); some cover only nonjournal material, for example, reports, patents, theses, and books. (This nonjournal material is, of course, often at least partially covered in the journal-oriented publications listed in Tables XIV.1-XIV.3. We shall discuss this coverage further in Section 3.2.) Table XIV.4 gives some illustrative examples of each of these categories.

Confronted by such a wide variety of secondary services, one is naturally curious to know how many of these are really used significantly by U.S. physicists. We have already noted, in Section 2.1 and Figure XIV.3, that U.S. physicists on the average devote only a very small amount of time to the use of secondary services, probably only a small fraction of an hour per week, but we have cautioned that the importance of this use may still be considerable. We shall comment at greater length in Sections 3.3 and 3.5 on how this time is divided among the various services. Here it will suffice merely to mention some figures obtained a few years ago by Slater and Keenan[21] on the numbers of U.S. physicists reporting the regular use of various abstract journals. About 29 percent of the reports of use named journals in Table XIV.1, with *Physics Abstracts* in the leading position; another 43 percent referred to journals in Table XIV.2, with *Nuclear Science Abstracts* leading and *Chemical Abstracts, Electrical Engineering Abstracts, Solid State Abstracts,* and *Engineering Index* closely spaced after it; nearly one sixth pertained to publications appearing in Table XIV.4, with *U.S. Government Research and Development Reports* in the lead and *Scientific and Technical Aerospace Reports* next, though considerably below it. It seems likely that if the various publications were compared on a basis of time devoted to them, the gap between the leaders and the others would be increased; however, no such statistics are available.

So far we have mentioned only those secondary services that supply printed lists of publications in designated subject areas. At least three other types of services should be mentioned; physicists have only recently begun to use these services, but they have considerable potential utility for the future. The most nearly established of these is the citation index, whose sole important representative for the physics community is the *Science Citation Index,* published by the Institute for Scientific Information in Philadelphia. (We shall discuss its utility in Section 3.4.) A newly available service, from which large employers of physicists can construct various local services to their employees, is abstract-index information on computer tapes; such tapes are now available from the American Institute of Physics with a very prompt coverage of the world's leading physics

TABLE XIV.2 Abstract Journals with Comprehensive Coverage of Journal Literature in Special Subfields of Physics or Other Areas Overlapping Physics[a]

Publication	Publisher	Scope	Physics Subfields[b]	Language	Entries (thousands)
Applied Mechanics Reviews	ASME, New York	Applied mechanics, etc.	Elasticity, plasticity, hydrodynamics, plasma flow, etc.	English	9.4 (1968)
Astronomischer Jahresbericht	Walter de Gruyter, Berlin	Astronomy	Astrophysics	German	13.5 (1968)
Bulletin of Thermodynamics and Thermochemistry	IUPAC, Univ. of Michigan, Ann Arbor	Thermodynamics and thermochemistry	Thermodynamic data	English	3.5 (1968)
Chemical Abstracts[c]	Chemical Abstracts Service	Chemistry and related engineering, biology, physics, etc.	Chemical, atomic and molecular and solid-state physics, etc.	English	233 (1968)
Electrical and Electronics Abstracts	IEE, London	Electrical and electronic engineering	Solid-state devices, electron physics, electromagnetic waves, dielectrics, some optics and plasma physics, etc.	English	30 (1968)
Engineering Index	Engineering Index, Inc., New York	Most areas of engineering	Most areas of applied physics	English	60 (1968)
Geophysical Abstracts	U.S. Geological Survey, Washington	Geophysics	Geophysics	English	7 (1968)
International Aerospace Abstracts	AIAA, New York	Aeronautics and astronautics	Aerodynamics, hydrodynamics, some geophysics, some instrumentation, plasmas, space physics, etc.	English	34 (1968)
Laser Abstracts	Lowry-Cocroft Abstracts, Evanston, Ill.	Lasers	Laser physics	English	0.15 (1968)
Mathematical Reviews[c]	Amer. Math. Soc., Providence, R.I.	Mathematics	Mathematical physics	English	15 (1968)
Metals Abstracts[c]	ASM, Metals Park, Ohio, and Institute of Metals, London	Metallurgy	Electronic and defect physics of metals	English	25 (1969)

astrophysical Abstracts	Boston	hydrology, geophysics, astrophysics	physics, earth and planetary physics		
Nuclear Magnetic Resonance Abstracts	Preston Technical Abstracts Co., Evanston, Ill.	Nuclear magnetic resonance	Nuclear magnetic resonance physics	English	0.18 (1968)
Nuclear Science Abstracts	U.S. Atomic Energy Commission, Oak Ridge	Nuclear science and technology	Nuclear physics	English	54 (1968)
Physics and Chemistry of Glasses	Soc. of Glass Technology, Sheffield	Physics and chemistry of glasses	Physics of glasses	English	0.5 (1968)
Physics in Medicine and Biology (Section "Abstracts")	Taylor and Francis, London	Physics in medicine and biology	Biophysics	English	1.3 (1968)
Referativnyi Zhurnal:	Acad. of Sciences U.S.S.R., Moscow				
Astronomiya		Astronomy	Astrophysics	Russian	7.2 (1965)
Geofizika		Geophysics	Geophysics	Russian	19.2 (1965)
Mekhanika		Applied mechanics	Mechanics, hydrodynamics, aerodynamics, plasma dynamics, elasticity, plasticity	Russian	21.6 (1965)
Yadernye Reaktory		Reactors and nuclear technology	Reactor physics	Russian	0.6 (1965)
Revue d'Optique Theoretique et Instrumentale	Institut d'Optique Theoretique et Appliquée and Syndicat Genéral de l'Optique et des Instruments de Précision, Paris	Optics and optical instruments	Optics	French	0.6 (1965)
Rheology Abstracts	British Society of Rheology, Pergamon Press, London	Rheology	Rheology	English	0.9 (1968)
Solid State Abstracts	Cambridge Communications Corp., Washington	Solid-state physics and devices	Solid-state physics	English	16 (1968)
Vacuum (Section "Classified Abstracts")	Pergamon Press, London	Vacuum science and technology	Vacuum and surface physics	English	1.6 (1968)

[a] Data are from Terry and Cooper[41] except where otherwise referenced.

[b] Roughly estimated from scanning subject headings.

[c] Not listed by Cooper and Terry.

TABLE XIV.3 Title Journals with Comprehensive Coverage of Journal Literature in Special Subfields of Physics or in Other Areas Overlapping Physics[a]

Publication	Publisher	Scope	Physics Subfields	Language	Entries (thousands)
Applied Science and Technology Index	H. W. Wilson Co., New York	Many technological fields	Applied physics	English	77 (1968)
Berichte der Deutschen Rheologischen Gesellschaft, E.V. (Section "Bibliographie, Documentation; Reports")	Deutsche Rheologische Gesellschaft, Berlin-Dahlem	Rheology	Rheology	German	3.5 (1963)
High Energy Physics Index	Gesellschaft für Kernforschung mbH, Frankfurt	High-energy physics	High-energy physics	English	9 (1968)
Journal of the Acoustical Society of America (Section "Current Publications on Acoustics")	AIP, New York	Acoustics	Acoustics	English	5.5 (1968)
Plasma Physics Index	Gesellschaft für Kernforschung mbH, Frankfurt	Plasma physics	Plasma physics	English	5 (1968)
Surface and Vacuum Physics Index	Gesellschaft für Kernforschung mbH, Frankfurt	Surface and vacuum physics	Surface and vacuum physics	English	–
Transatom Bulletin	Commission of the European Communities, Kirchberg, Luxemburg	Atomic energy	Some nuclear	English	10 (1968)

[a] Data are from Terry and Cooper.[41]

TABLE XIV.4 Examples of Types of Secondary Services Other Than Those Listed in Tables XIV.1-XIV.3

Type	Publication	Publisher	Material Listed	Language
Limited or selective coverage of journal articles	*Notizario* ("Sommario" section)	Com. Naz. En. Nuc.	Nuclear science and technology	Italian
	Le Vide: Technique, Applications ("Documentation" section)	Soc. Franc. Ing. et Tech. du Vice	Vacuum technology	French
	Journal of the Acoustical Society of Japan (abstract section)	Acoustical Soc. of Japan	Acoustics	Japanese
Coverage only of reports, dissertations, patents, etc.	*List of References on Nuclear Energy*	IAEA	Reports on nuclear energy	English
	Scientific and Technical Aerospace Reports	NASA	Reports on aeronautics, astronautics, etc.	English
Book listings	*Quarterly Checklist of Physics* (including astronomy and astrophysics)	American Bibliographic Service	Books, brochures, etc. in these fields	English
	Physikalische Blätter (section "Bücher")	Physik-Verlag GmbH	Physics books	German
Regional coverage	*Physics Express*	Internat. Physics Index, Inc.	Physics articles from U.S.S.R.	English
	U.S.S.R. Scientific Abstracts: Physics and Mathematics	U.S. Dept. of Commerce	Physics and mathematics articles from U.S.S.R.	English
	Indian Science Abstracts	Indian National Scientific Documentation Center	Indian scientific and technical literature	English
	Dissertation Abstracts	University Microfilms, Inc.	U.S. dissertations	English
	U.S. Government Research and Development Reports	U.S. Dept. of Commerce	Government reports on science and technology	English

journals,[42] from the Institution of Electrical Engineers, with the somewhat slower but more comprehensive information of *Physics Abstracts* and their other two abstract journals, and from Chemical Abstracts Service. Some services are also available that supply an individual scientist with currently published papers that match an individual interest profile that he can select. The two most widely available such services of interest to physicists are the Mathematical Offprint Service of the American Mathematical Society (limited, of course, to areas that could be called mathematics) and the Automatic Subject Citation Alert of the Institute for Scientific Information. The latter provides lists of current papers that match certain specifications selected by the subscriber in terms of authors, citations, institutions, and the like. The papers themselves can also be supplied if desired. In addition to these widely available services, certain organizations provide various types of selective-dissemination services to their employees; these have been reviewed, though without specific reference to physics, by Connor[43] and East.[44]

3.2 BULK AND COVERAGE

We have already given, in Tables XIV.1–XIV.3, some reasonably current figures on the numbers of items covered by the leading abstracting and indexing publications and a few words on the ground they purport to cover. We shall now take a more detailed look at these topics.

Figure XIV.10 compares the growth over the years in the numbers of entries in *Physics Abstracts, Physikalische Berichte,* and *Referativnyi Zhurnal, Fizika.* Besides illustrating the well-known exponential growth of the literature from decade to decade, punctuated by a dip in the war years, the curves show some rather wild fluctuations that have come at different times for the different journals. These have undoubtedly been due to fluctuations in the finances of the journals and to administrative decisions affecting the breadth of subject coverage, the number of journals scanned, and the inclusion of reports, patents, and other nonjournal material. The principal lesson to be learned from Figure XIV.10 is that one must proceed with great caution if one wishes to compare the growth of the literature of physics with that of, say, chemistry, by comparing the sizes of *Physics Abstracts* and *Chemical Abstracts.* (*Chemical Abstracts* represents an extreme case, in that quite a large proportion of its entries are items that are clearly in the domains of other well-recognized disciplines, such as biology or physics, and only peripherally related to chemistry.)

It is interesting to note that despite the fluctuations evident in Figure XIV.10, the number of journals scanned for *Physics Abstracts* has increased linearly with time, from 200 in 1920 to over 800 in 1968.[38]

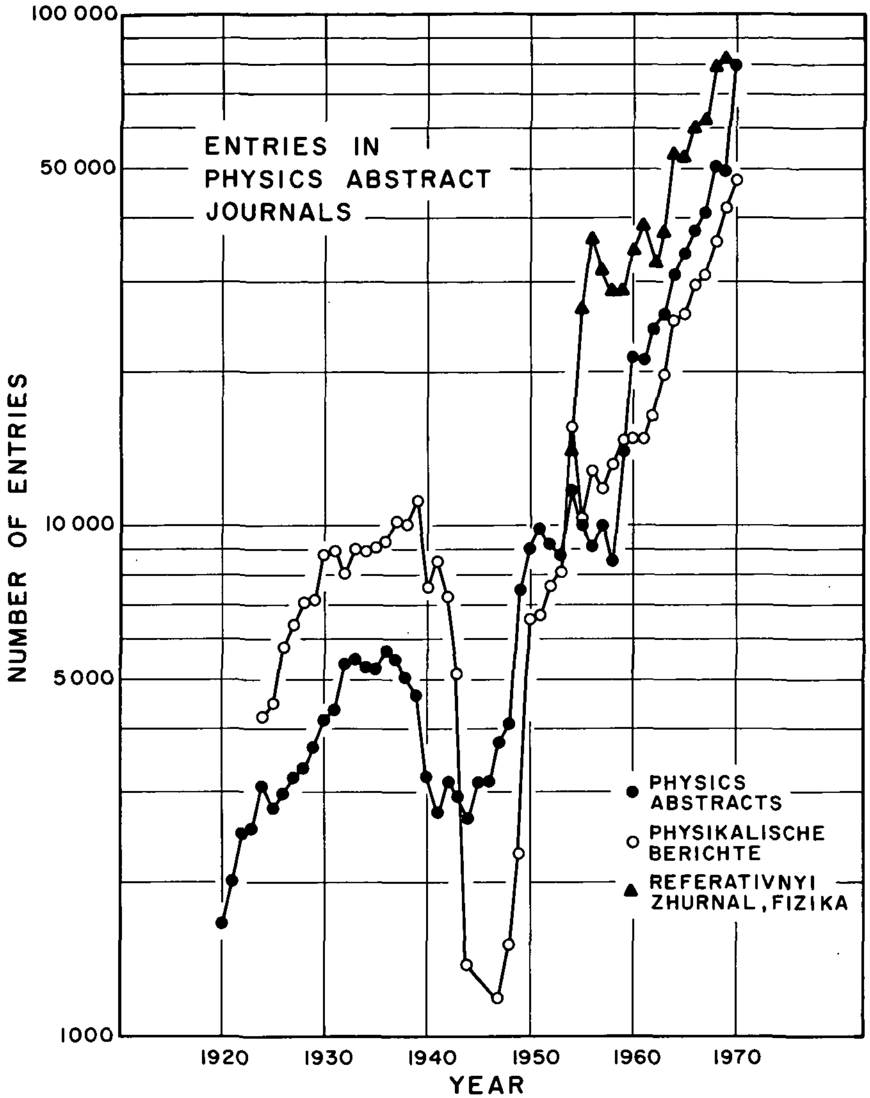


FIGURE XIV.10 Growth in time, of total entries in three abstracting journals covering the literature of physics.

It would obviously be of considerable interest to compare abstract journals in regard to their relative numbers of entries for journal articles, reports, theses, patents, books, and articles in books and to follow the evolution in time of all these quantities. Unfortunately, very few data of this sort are available. Figure XIV.11 shows a few scattered examples; some older data, considerably more detailed, are given in ICSU-AB,[45] Table 31, and ICSU-AB,[46] Table 45. Clearly, different abstract journals offer quite different mixes of these elements.

Another interesting, though tedious, comparison one can make among abstract journals has to do with the thoroughness of their coverage of various classes of material. Again, data of this sort are scarce; the most extensive sources are a comparison of *Physics Abstracts* with *Nuclear Science Abstracts*[47] and the ICSU-AB studies.[46,47] Figure XIV.12 shows a few available examples from these and other sources. From these examples and from Figure XIV.11 we can draw several very tentative conclusions:

1. All leading abstract services in physics (top half of Table XIV.1) at present probably cover the really significant basic-research journal literature pretty thoroughly. For example, the gap in coverage of solid-state physics by *Physics Abstracts,* which was conspicuous a dozen years ago[48] seems to have been pretty well filled.
2. No abstract journal has at all complete coverage of articles marginally identifiable as significant research, of articles in all areas peripheral to physics, or of reports and other nonserial literature.
3. *Physics Abstracts* covers journal literature published in the Soviet Union fairly well (85 percent or so), as judged by comparison with *Referativnyi Zhurnal*, but is probably somewhat more deficient in its coverage of literature published in other parts of Eastern Europe. There are modest differences in the distribution of coverage between *Physics Abstracts* and *Bulletin Signaletique* (ICSU-AB,[45,46] Tables 11–14).
4. Abstract journals specializing in a particular subfield sometimes cover this subfield more completely than those that try to cover all of physics (for example, *Nuclear Science Abstracts* versus *Physics Abstracts*) but sometimes do not.

To underscore point 2, it is sobering to note the very large number of journals that occasionally contain an article of relevance to physics research. This finding is illustrated by Figure XIV.13, taken from Keenan and Brickwedde,[49] which shows the way in which the yield of articles abstracted in *Physics Abstracts* 1965 from the n most productive journals varied with n. From this and some of the evidence in Figure XIV.12, one can guess that scanning several thousand journals for *Physical Abstracts* might well increase the yield of articles deemed relevant for physics by several percent. *Production of an abstract journal must always be a compromise between the desire for completeness and the desire for economy.* In addition, it is

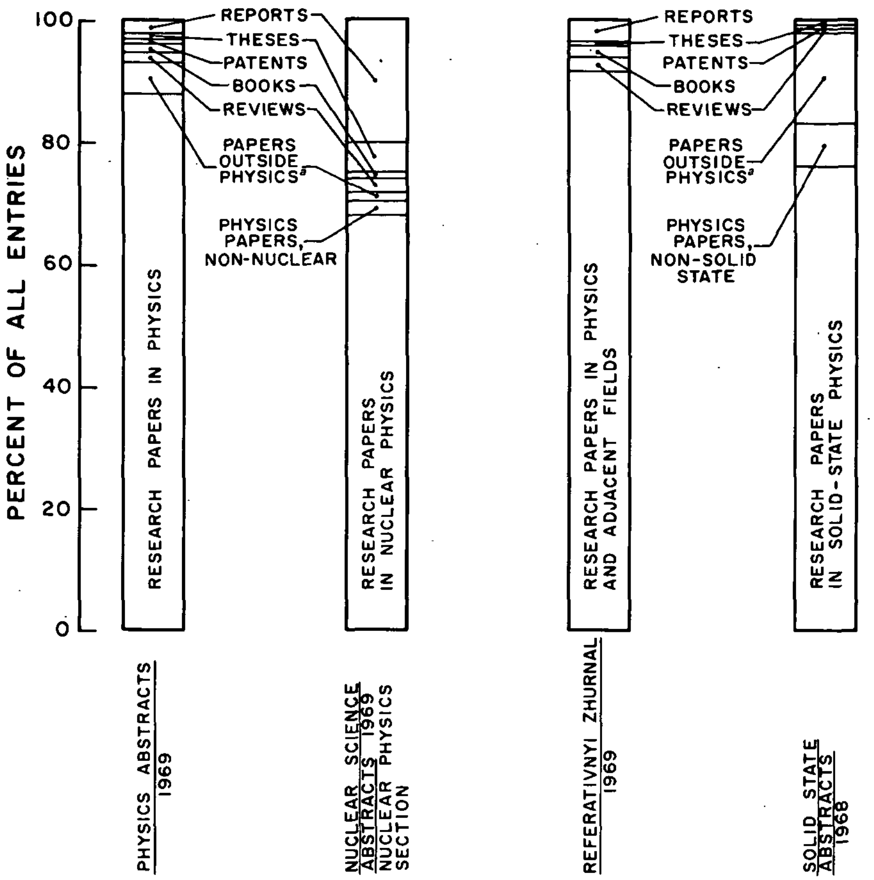


FIGURE XIV.11 Rough breakdown of the entries of several abstract journals by type of material. Since the definitions of the different categories of entries overlap, the order of decisions must be specified to clarify the meaning of the numbers. First, each entry was classified according to whether it was "published" (i.e., available, other than as an abstract, in a regular journal or in a book sold on the open market). Published material was classified into research papers, reviews or compilations, books, popularizations, and patents. Unpublished material was classified into theses and reports, including individual-paper components of reports. Research papers are defined as published (as just indicated) articles or letters reporting new research results.

[a]The boundary between physics and nonphysics was chosen rather arbitrarily, perhaps encroaching a little more on chemistry, geophysics, and like sciences than an impartial boundary would; however, it is hoped that the boundary has been drawn consistently for the different abstract journals.

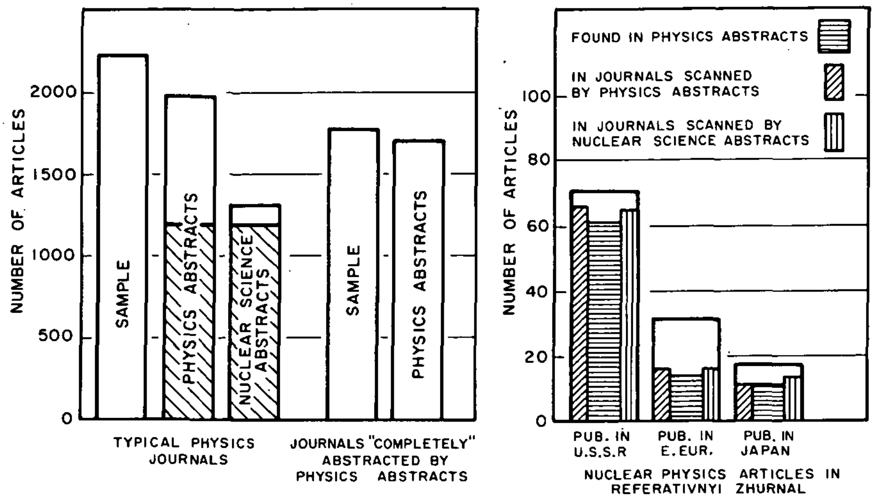


FIGURE XIV.12 *Left:* Results of a search (Greer and Atherton 1966) for 2246 articles from 29 typical physics journals of the world. The second and third bars show how many of these were found in *Physics Abstracts* and *Nuclear Science Abstracts* through the middle of 1965. The shaded portions indicate articles found in both. The next pair of bars shows the number of these articles that were in articles "covered completely" by *Physics Abstracts* and the number actually found (a possible measure of the undercounting due to long time lags). *Right:* Results of a search in several years of *Physics Abstracts* for nuclear physics articles from three regions abstracted in *Referativnyi Zhurnal* 1965 and 1969 and comparison of the media in which they appeared with the lists of journals scanned by *Physics Abstracts* and *Nuclear Science Abstracts.* (Many of the Eastern European articles were in conference books, not journals.)

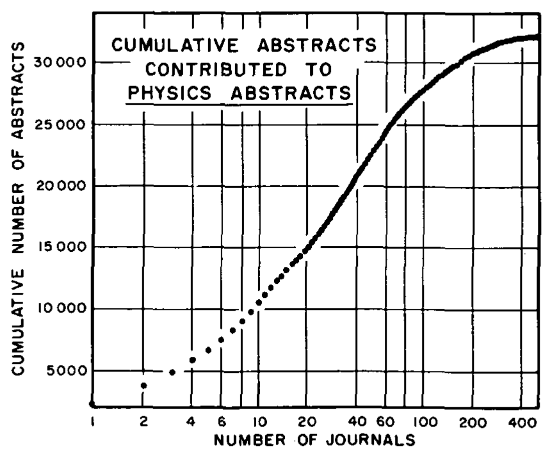


FIGURE XIV.13 Cumulative number of abstracts contributed to *Physics Abstracts* 1965 by the *n* most prolific journals, as a function of *n*. [Source: Keenan and Brickwedde[49].]

often desirable to restrict coverage to prevent the bulk of the journal from becoming a hindrance to users.

So far we have discussed only abstract journals. Title listings can be dealt with more briefly. Current sizes have already been given in Tables XIV.1 and XIV.3; *Current Papers in Physics* coincides in coverage with *Physics Abstracts,* while *Current Physics Titles* covers, roughly, the more important two thirds of these papers; *Current Contents, Physical and Chemical Sciences* (1971) gives complete contents of over 800 journals, which include about one fourth of those scanned (but in general far from completely covered) by *Physics Abstracts.* *

There has been little study of the coverage of reports, patents, theses, books, and the like in both standard abstract and title journals and special secondary publications, so there is little we can say here other than to point out the obvious; for example, for reports such coverage can never be as complete as that for journal literature and it varies widely from one secondary publication to another; book listings in the most commonly used physics abstract journals are rather unsatisfactory; tape services reflect the coverage of the associated abstract services.

3.3 TIME LAGS, AVAILABILITY, AND OTHER CONSIDERATIONS

The utility of the various secondary services to scientists depends on two types of factors. One is the amount of information the services can supply: This involves both their coverage, already discussed in Section 3.2, and the nature of the entries and their indexing, which we shall take up in Section 3.4. The other type of factor, which we shall discuss now, has to do with availability, both spatial and temporal.

The vast expansion of the literature in the last two decades has had a drastic effect on the pattern of circulation of the main abstract journals; these have become enormously more expensive to produce, and rising subscription prices and inconvenience of storage have caused subscriptions by individual physicists, once a sizable fraction of the total, to decline to a negligible number. Figure XIV.14, for example, shows the pattern of prices and subscriptions for *Physics Abstracts* over the last decade.

The lack of an observable effect of recent sizable price increases on institutional subscriptions contrasts strikingly with the strong effect of these and earlier increases on individual subscriptions. But it would be unwise to conclude that institutional purchases are unaffected by price. The more

* Note that, according to what has been said in the preceding paragraph, this latter fact may not imply seriously reduced coverage of physics; moreover, some of the missing journals are covered in *Current Contents, Engineering and Technology.*

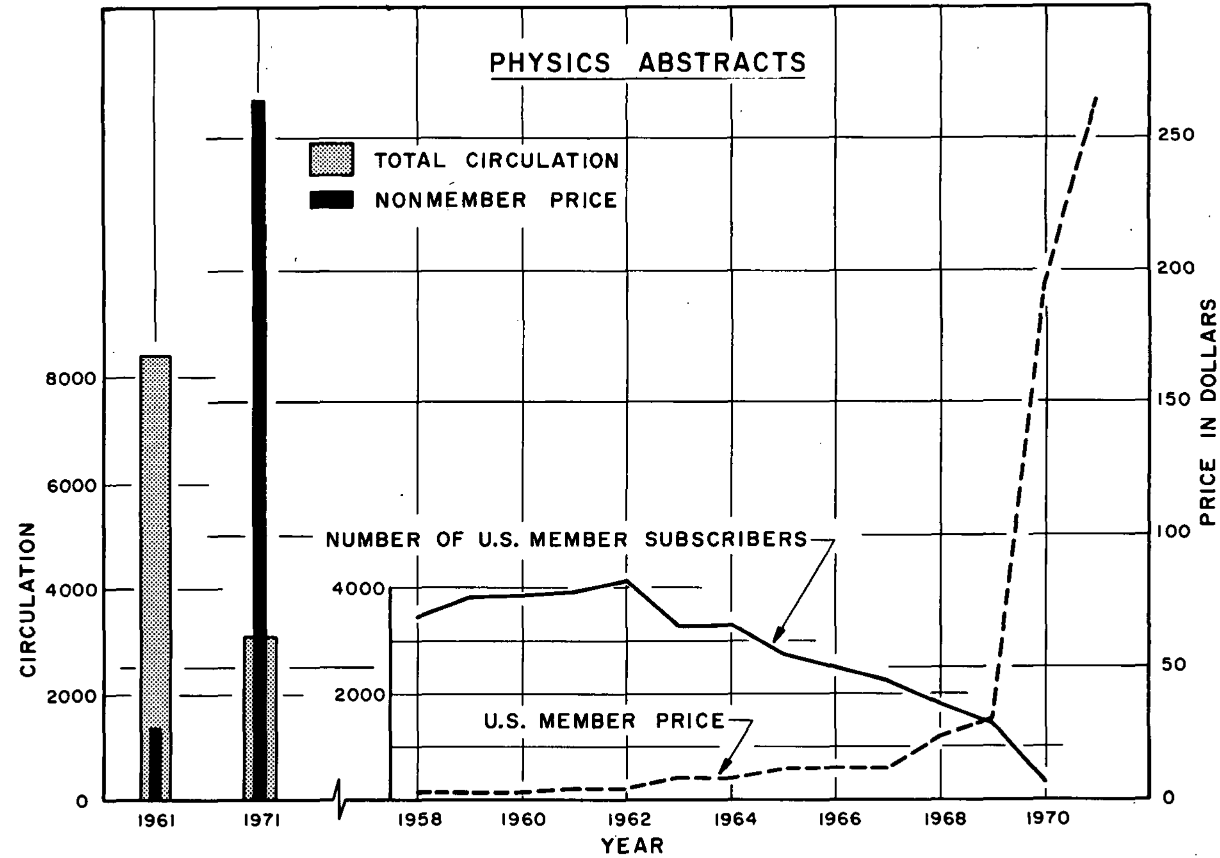


FIGURE XIV.14 Changes in price and circulation for *Physics Abstracts* with time. *Left:* Total circulation (shaded bars, scale at left) and nonmember price (black bars, scale at right) for 1961 and 1971. *Right:* Number of AIP member subscribers (solid curve, scale at left) and price charged to these (dashed curve, scale at right). Subsidy was removed in 1970.

useful abstract journals are often available in duplicate copies in several branch libraries of large institutions, and when they get too expensive some of these duplications may be eliminated, though there will doubtless be a strong initial tendency to continue existing subscriptions.

A plot of circulation against unit price, as shown in Figure XIV.15 for a number of abstract journals in fields in or bordering on physics, is illuminating. There is quite a noticeable inverse correlation of circulation with cost, modified by such obvious factors as, for example, potential user population (large for *Chemical Abstracts*), quality of abstracting (high for *Applied Mechanics Reviews* and *Mathematical Reviews*), and availability to society members at a price below that shown (*Applied Mechanics Reviews* and *Mathematical Reviews*). It would be quite consistent with the figure to speculate that as one passes from the price of *Nuclear Science Abstracts* to that of *Physics Abstracts* one loses most individual subscribers and some of the multiple institutional subscriptions, and that another factor of three or so in price would cause a sizable loss of institutional subscriptions. Although the rapid rise in both price and bulk with time, shown in Figures XIV.14 and XIV.10, suggests that an abstract journal might try to retain its individual subscribers by subdividing, the experience of smaller and more specialized abstract journals does not hold much promise in this direction: The inverse correlation of circulation with price per abstract, shown in Figure XIV.15, is not favorably modified for the smaller publications.

It can hardly be doubted that the decline of widespread individual subscriptions to abstract journals and the probable decrease in the availability of multiple institutional subscriptions (for example, subscriptions for use in sublibraries), have decreased the frequency with which physicists use these journals. Fortunately, a part of this decrease has been compensated by the growing availability of title listings, which in many cases (for example, *Current Papers in Physics* and *Current Physics Titles*) have prices sufficiently low to attract sizable numbers of individual subscribers. However, title listings are used mainly for current awareness, and most of them are not indexed in a way that makes them suitable for retrospective searching, even if the user were willing to forgo having an abstract. Thus the present pattern undoubtedly represents a net loss in the usefulness of secondary services, as compared with earlier days when abstract journals were more plentifully distributed. As we have noted earlier, in Section 2.1, many studies have shown that the use of any communication medium, be it a library volume or a talkative friend, decreases extremely rapidly as its distance from the user's office or laboratory increases.

Thus it is not unlikely that the benefit the physics community receives from abstract journals could be appreciably increased if more copies could be made available in sublibraries and the like. (The bulk is now so great

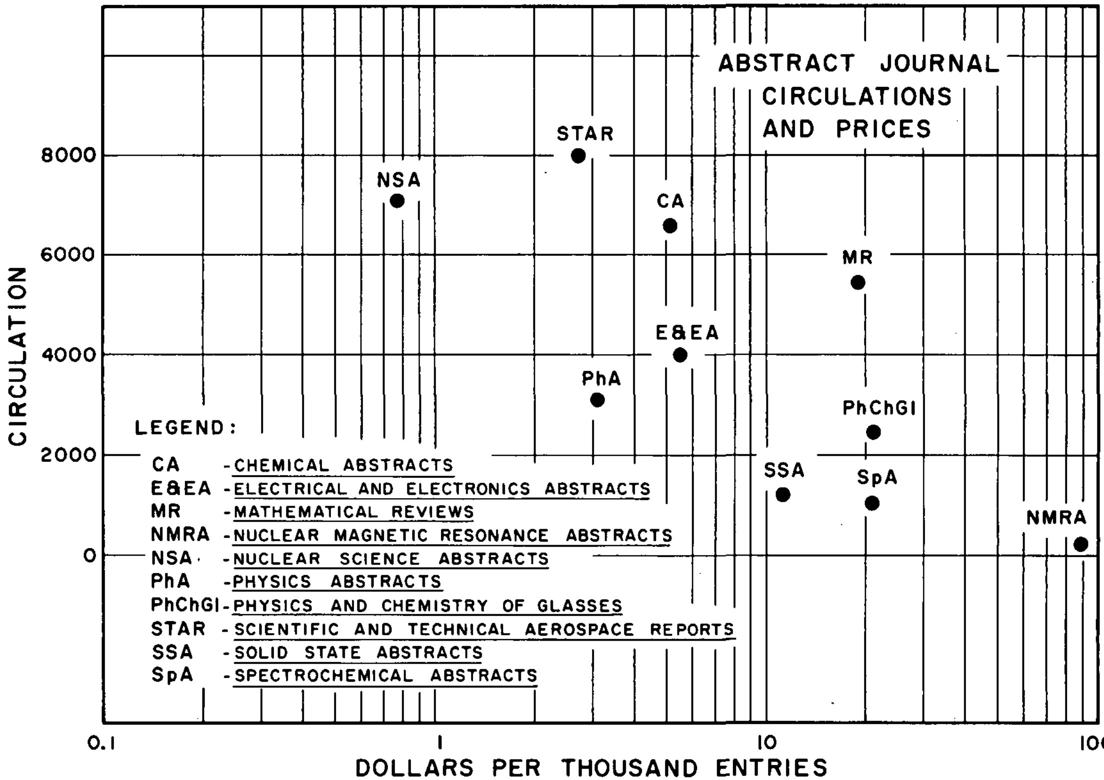


FIGURE XIV.15 Circulations and prices per abstract for abstract journals in physics and related fields. Data are taken largely from Ulrich's *International Periodicals Directory* (1968–1969 or 1971–1972).

that probably few individual physicists would wish to own them even if they were free.) In this respect, *Nuclear Science Abstracts*, subsidized by the U.S. Atomic Energy Commission and sold at a low price, probably makes better use of the resources that go into its production than does *Physics Abstracts*, which undertakes to cover its whole production cost by selling subscriptions at a flat rate ($264.00 in 1971) and which therefore reaches a smaller proportion of its potential subscribers. We shall return in Section 3.6 to the question of how the benefit of distributing additional copies of abstract journals might compare with the cost of producing the extra copies and to the question of whether a subsidy or alternative pricing policy should be considered.

The currency of an abstracting or indexing publication is a significant factor in its utility, especially if it is to be used for current awareness. Fortunately, great improvements have occurred in this area in the past decade, though there is wide variation from one secondary publication to another. (There are also interesting geographical correlations: See, for example, ICSU-AB,[45] Table 41B; and ICSU-AB,[46] Tables 30–32.) Whereas about ten years ago the median time lag between appearance of a paper and the appearance of its abstract in *Physics Abstracts* was of the order of five to eight months and delays of well over a year were not uncommon (some services, for example, *Referativnyi Zhurnal*, were much worse), today the median time lag of *Physics Abstracts* is only about 3.5 months, and delays of as much as a year are very rare. *Referativnyi Zhurnal* now has a five- to six-month median lag. Much of the improvement for *Physics Abstracts* is due to communication of abstract information by journals in the proof stage, before the journals are issued; the information on all AIP publications, for example, is sent by air to *Physics Abstracts* at this stage. The possibilities for use of computer-tape records are obvious and are being actively pursued. Cooperative agreements, not only of the sort mentioned between primary and secondary services but also between different secondary services, especially those in different countries, are becoming increasingly necessary and important. (We shall discuss the outlook for the future in Section 3.6.)

Title listings have become important in physics only in the last five years or so, although certain fields (for example, chemistry) have used them for a longer time and some research organizations have had services of this sort for internal use for many years. Since these services are directed mainly toward satisfying current-awareness needs, promptness is especially important for them. Thanks to the techniques mentioned in the preceding paragraph, very fast listings are now being achieved, especially when, as with the AIP publication *Current Physics Titles*, comprehensive coverage of minor journals is not attempted. Thus, for example, *Current Papers in*

Physics (the title journal corresponding to *Physics Abstracts*) publishes classified title lists with a median time lag of about 2.5 months.

3.4 CONTENT, ORGANIZATION, AND INDEXING

Secondary services can provide a variety of types of information. Consider first the individual entries themselves. If these are abstracts, they can vary from "indicative" (merely indicating what topics are addressed in the paper being abstracted) through "informative" (summarizing the principal conclusions of the paper) to "critical" (giving the abstractor's evaluation of the paper). The swelling volume of the literature has driven many abstract journals to rely mainly or even entirely on authors' abstracts, hence abandoning critical abstracts entirely and abandoning any control over whether the abstracts published are informative, indicative, inadequate, or even invisible. Among the abstract journals listed in Table XIV.1, *Physics Abstracts* relies almost completely on the authors' abstracts appearing with the original publications; except for a very few cases of material published in foreign languages, no abstract is supplied if the author did not publish one. This policy has evolved gradually: In 1964, a sizable minority of the abstracts were prepared by abstractors.[46] *Referativnyi Zhurnal* is at the other extreme, using authors' abstracts for only a minor fraction of the entries and signed abstractors for the rest. *Physikalische Berichte,* too, has a majority of its abstracts prepared by signed abstractors, though authors' abstracts are used more frequently than in *Referativnyi Zhurnal. Bulletin Signaletique* seems to use mainly authors' abstracts. In general, abstract journals that have been started fairly recently (for example, *Nuclear Science Abstracts, Solid State Abstracts*) tend to rely heavily on authors' abstracts.

As one would expect, authors' abstracts, though usually of the informative type, are often only indicative, or even inadequate, while those prepared by abstractors are (in these fields) nearly always informative. Occasionally, they may verge on being critical evaluations, but usually the abstract journals make no concerted effort in this direction; *Applied Mechanics Reviews* and *Mathematical Reviews* are notable exceptions.

Various types of entries are also possible for title listings: Although in most cases only the original title is given, this is sometimes augmented with an indication of whether the paper is experimental, theoretical, or both, with keywords relevant to its content, or with subject index numbers. The new AIP journal, *Current Physics Titles*, is experimenting with additions of this sort.

Besides listing the authors' names and the bibliographic references, abstract and title journals are increasingly adopting the practice of listing authors' institutions. *Physics Abstracts* started doing this in 1969, *Physikalische Berichte* in 1937, and *Nuclear Science Abstracts* at its inception in 1960. However, this information can be given only when it was given in the original paper.

Arrangement and indexing are all-important for the utility of any secondary service; almost no one skims through an abstract or title journal cover to cover. Nearly all such journals arrange their entries according to a subject classification scheme, and most but not all supplement classification with duplicate entries or cross references for items that could logically appear in more than one place. In *Physics Abstracts,* for example, the number of such cross references is about 30 percent of the (unduplicated) total number of entries, and the great majority of the subject headings at the finest level have less than ten entries per issue. Some rough data for a few other abstract journals are summarized in Figure XIV.16. Although one can see the expected tendency toward more frequent cross referencing as the

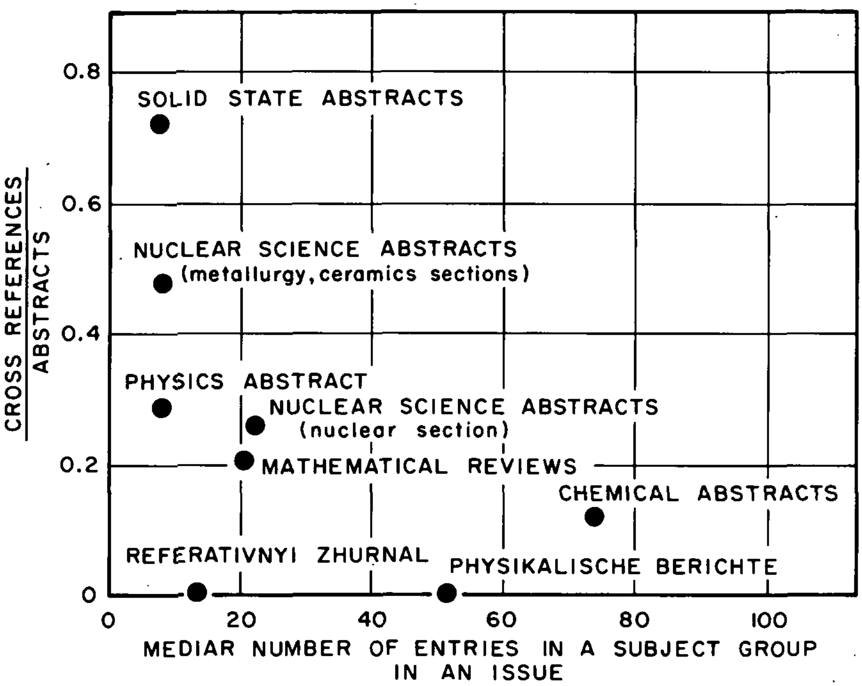


FIGURE XIV.16 Distribution of several abstract journals with respect to the fraction of abstracts that are cross referenced from one subject category to another and the median number of entries, per issue, in the finest subdivisions of the subject grouping.

subject groupings get smaller, the fluctuations from journal to journal are large, some having no cross referencing of individual papers at all.

Subject and author indexes, cumulated at annual or other intervals, are of course the most important determinants of accessibility. Table XIV.5 shows some features of these for the abstract journals of Table XIV.1 and a few others. Author indexing is a fairly straightforward task, and the various refinements that may assist the user (listing of full titles of papers, listing under every author of a multiauthor paper, and the like) can be introduced or withdrawn independently of one another. Subject indexing is more involved. Not only is it necessary, as it already is for the subject arrangement in each issue, to have relevant subjects assigned by someone with expertise, but there is also an intrinsic conflict between the two desiderata of making the meanings of the subject categories unambiguous and keeping down the length of the list to be scanned for each category. Fine subdivision (many subject categories) achieves the goal of reducing the length of listings but makes the classification of a given paper under a given heading more ambiguous. Listing the paper under every subject that could reasonably be identified with it will relieve this problem and ensure that the user will not miss it, but at the cost of making the number of entries under each subject larger again. The publishers of abstract journals thus have some difficult decisions to make.

Our discussion of the information content of secondary services would be glaringly deficient if it did not point out the extremely rich information content of citation indexes. The author of a scientific paper reveals a great deal about its intellectual content by the material that he chooses to cite. Although the set of citations will normally not reveal the nature of the new contributions made by the citing person—as an abstract does—it will usually reveal a great deal about the subject matter and methodology. The great advantage of citations is their convenience as an access tool to tap a rather sizable pool of information based on the expertise of the authors of papers. The simplest (though by no means the only—see Section 3.6) use of this resource is to look for papers that cite at least one of several standard papers on some topic one is interested in; one will then retrieve most of the new papers on this topic, albeit with a fair number of unwanted ones. Despite the latter "noise," we shall see, in Section 3.5, that citation index searching compares favorably with other methods in certain fields.

3.5 PATTERNS OF USE OF SECONDARY SERVICES

We have already seen, in Section 2.1 and Figures XIV.3 and XIV.4, that physicists spend only a small fraction of their time using secondary services and get only a small fraction of their leads to useful information from these; at the same time, we have seen in Figure XIV.5 that a sizable number of physicists, at least, consider abstract and current-awareness journals to be important media for keeping up to date in their fields. Figure XIV.4 also seems to indicate that in the physics community more items of useful information are obtained from browsing in these media (the current-awareness use) than from retrospective searching. A further study[50] of atomic and molecular physicists has shown their use of abstract journals to be of the order indicated in Figure XIV.3, with *Physics Abstracts* well ahead of all others.

In all these respects, physicists differ markedly from scientists in some other fields, notably chemistry. According to the penultimate column of Figure XIV.3 and note *h*, industrial research chemists spend over half as much time with secondary services as they do with primary literature, and about half of this time is spent in retrospective searching. Much of chemists' searching is done on particular substances—witness the great attention given to substance and formula indexing in *Chemical Abstracts*—whereas physicists, though sometimes interested in specific substances or nuclear species, usually seek information that is much less easy to identify by an index label. They therefore find the subject indexes now available much less rewarding than chemists do, and hence use them less.

This last point is illuminated a little by two studies of potential search requests from physicists. In 1961 and 1962, the American Institute of Physics undertook, as part of an effort to design improved literature access tools, to find out the nature of the literature-searching requests that might be generated by nuclear physicists if an "ideal" searching service could be provided.[51] The great majority of the search requests were for information on particular nuclear species or reactions. But a fair number were of other types such as the following:

"All optical-model fits to low-energy elastic scattering data"
"Theoretical and experimental work on collective excitations of highly deformed nuclei in the rare-earth and heavy-element regions"

All the requests in this study were of the dragnet type, that is, the physicists were not asked to make requests for isolated pieces of information. In the other study (Herring, 1966, unpublished data), however, a group of solid-state physicists were encouraged to ask any questions, answers to which were needed by them and were considered likely to be available somewhere in the literature. Only a minority of these turned out to be describable in terms of standard properties of standard materials (perhaps because the physicists did not feel in need of help for such questions).

TABLE XIV.5 Characteristics of Indexes of Typical Abstract Journals

	Subject Index			Author Index			
Abstract Journal	Type[a]	Average Entries per Abstract[b]	Average Entries per Subject Heading[c]	Treatment of Title	Treatment of Multiple Authors[d]	Other Indexes	Cumulations
Physics Abstracts (1969)	Alphabetical with subcategories	2.8[e,f]	41	Abbrev.	Each	Book, bibliography, patent, conference, report	6 mo
Physikalische Berichte (1969)	Hierarchical	1.5	53	Single-word	Each		Annual
Referativnyi Zhurnal, Fizika (1969)	Alphabetical	1.1	42	None	Each		Annual
Bulletin Signaletique (1968)	Alphabetical with subcategories	1.4	Not easily comparable	None	Each		Annual
Nuclear Science Abstracts (1970)	Alphabetical	2.9[e,f]	20	Full	Each	Corporate	Annual, 5 yr

Solid State Abstracts (1968)	Alphabetical with subcategories	3.3	Few[c]	None	Each	Journal sources	3 mo
Chemical Abstracts	Alphabetical with subcategories	8[e, f]	Few[c]	Full	First, with cross refs.	Formula, patent	6 mo, 5 yr
Mathematical Reviews	None			Full	Each		Annual, 5 yr, 20 yr
Applied Mechanics Reviews	Alphabetical with subcategories	1.6	7	Full	First, with cross refs.		Annual

[a] Alphabetical means an alphabetical listing of keyword topics; hierarchical means an arrangement according to the elements of a logical classification scheme. With subcategories means that under each alphabetical main heading further specializations are listed, again alphabetically.

[b] Average number of places in which a given abstract appears in the subject index.

[c] Average number of entries in the annual, semiannual, or quarterly index (see last column) that a reader must scan under a single subject heading (finest subdivision). (Occasionally, however, as occurs for *Chemical Abstracts*, an alphabetical arrangement of substances under a property heading can reduce the figure given to unity or near it.)

[d] Each means that the identification number and title information for a given abstract can be found under the name of any one of its authors. First, with cross references means that this information can be found only under the name of the first author, but that there are entries under the names of other authors, referring to the first author.

[e] Two keyword listings, one for "topics," the other for "materials."

[f] Rather similar figures were obtained in 1959: *Physics Abstracts*, 2.4 entries per abstract, *Nuclear Science Abstracts*, 3.0 entries per abstract, and *Chemical Abstracts*, 6 entries per abstract.

Typical questions of other types were:

"Has anyone ever tabulated the crystal-group analogs of the $6-j$ symbols?"

"What is the experimental evidence for the superfluidity of monolayer helium films?"

"Is there a more complete treatment of the relation of the two gyromagnetic ratios (the Zeeman ratio g and the gyromechanical ratio g') than the old work of Kittel and Van Vleck?"

"What is known about the mechanism of production of a hole in a solid by a laser beam insufficient in intensity and duration to produce the effect by heat alone?"

Although one could make some progress on most such questions using existing subject indexes, the search generally would be very laborious, and few physicists are willing to undertake it.

The problems posed for the designers of indexes by the conceptual rather than taxonomic nature of most work in physics is nicely illustrated by an unpublished study done by Maizell in 1960 of speed of retrieval. A sample of graduate students were given abstracts of *Physical Review* papers, without the authors' names or the journal citation. Some were asked to locate the articles using one subject index, some using another index. It was found that retrieval was significantly slower with the more extensive subject indexes: For example, that of *Chemical Abstracts* required, on the average, nearly twice as much time as that of *Physics Abstracts.*

The utility of citation indexes for retrospective searching is not yet as widely appreciated as it should be. Some of the sample questions given above could obviously be pursued with profit in this way. The documentation literature contains a few more quantitative studies in nonphysics fields. For example, Spencer (1967) compared searching via *Index Medicus* and *Chemical Abstracts* with searching via *Science Citation Index* with respect to retrieval rates for relevant references on the properties of thalidomide (a subject well suited to the indexing schemes of the former journals). She found that in the first few hours of work, searching in the citation index, using a reasonable strategy, gave a significantly larger yield of relevant references than searching in the conventional media; in later stages of the search, however, the conventional media proved more helpful.

Let us now turn to the use of current-awareness publications. The most extensive survey of the use of these was made in 1965 as part of a study to evaluate the extent of the need for the new publication *Current Papers in Physics* that was just being started and to provide clues regarding features that would be useful in this publication.[21,52] In this study, which we have already cited briefly in Section 2.1 and Figure XIV.5, a questionnaire on current-awareness habits was sent to large and representative samples of U.S. and U.K. physicists; unfortunately, the significance of the results is clouded by the low percentages of usable returns (30 percent in the United Kingdom, 26 percent in the United States). With this caution in mind, we note some of the findings:

1. Only a small minority of the respondents (22 percent in both samples) at that time relied on current-awareness journals for keeping up with their field; the "figure of merit" for such publications, as presented in Figure XIV.5, was far below that of journal scanning, scanning of abstract journals, and personal contacts.
2. A rather larger number of respondents did mention, however, that they looked regularly at some current-awareness journal. For the U.S. sample, the one most often mentioned (19 percent of the respondents) was *Current Contents, Physical Science.*
3. An even larger proportion of the respondents mentioned using other current-awareness services, most often accessions lists of their own organization. (These might often have been only book listings.)

Unfortunately, no comparable study of the extent of present-day use of *Current Papers in Physics* seems to have been made, though an extensive study of the *nature* of its use has been published.[52-54] The available evidence suggests that it is still not widely used, despite an extensive promotional campaign that reached essentially all U.S. physicists. The circulation of this journal, currently about 1700, is considerably below that of *Physics Abstracts* (Figure XIV.14). Clearly, the low price ($14.50 to members of AIP societies) has not attracted many individual subscribers; from what is known generally about the decrease of use of information resources with increasing distance from the user's office, it seems unlikely that there is widespread regular use of library copies. Small-scale surveys we have made seem to confirm that the great majority of physicists do not regularly use either *Current Papers in Physics* or any other current-awareness publication of broad coverage. Much smaller still, though no figures are publicly available, is the use of the computerized personal alerting service offered by the Institute of Scientific Information.

Is this apparent low rate of use of titles journals due to simple inertia and lack of awareness of their utility, or does it imply that time invested in scanning them is, in fact, unproductive? No definitive answer is known, but there is some evidence that the former alternative is the true one. For example, a sizable majority of physicists at the Bell Laboratories regularly uses an internal titles listing fairly similar in coverage, layout, and promptness to *Current Papers in Physics* that has been available for many years. It is hard to see why physicists in other organizations should not find *Current Papers in Physics* equally useful. We thus conclude tentatively:

4. The use of titles journals and other current-awareness services is currently probably far below what would be optimum for the physics community.

It will be interesting to see whether AIP's new sectionalized journals, *Current Physics Titles,* have more appeal.

3.6 TECHNOLOGY, ECONOMICS, AND OUTLOOK FOR THE FUTURE

What do all the data we have been examining imply for the policies of scientific societies, governmental agencies, and research and educational institutions? Such policies should be determined, of course, by balancing costs against benefits for various gradations of possible secondary services. But one must resist the temptation to make the comparison in too narrow a context, for example, in terms of what a single producing organization can do without subsidy or cooperative agreements with other organizations. We urge the following procedure: *Estimate the total benefit to society's scientific enterprise that will result from the provision of a hypothetical secondary service, or modification of one, using methods similar to those discussed in Sections 2.4 and 4.8. Estimate the cost, to society as a whole, of providing the proposed service. If the benefit exceeds the cost, try to work out a method of financing the service that will be feasible, stable, and conducive to efficiency.*

The issues involved are not peculiar to physics, and they almost always transcend national boundaries. They have, in fact, received a great deal of interdisciplinary and international attention.[2,55] Here we shall merely offer a few comments from our observations in physics; insofar as physics is a typical field, some of these comments will be relevant to the problems of secondary services in general. A few of our comments, on the other hand, will point to ways in which physics differs from many other fields.

Let us start by considering production costs. For an abstract journal, a part of these costs consists of components of the same sort as are encountered in the production of a primary journal (copy editing, composition, paper, presswork, mailing). (See also Section 4.2.) But there are differences. Instead of the editorial criticism, refereeing, and alteration of manuscripts, there are the scanning of journals, selection of articles to be covered, and, if necessary, translation of abstracts from a foreign language or preparation of abstracts by an abstractor. The average cost of these operations is considerably less, of course, than that of the scientific editing of a paper, but their cost per printed word is considerably more. Information collected on the production of *Physics Abstracts* over a number of years, combined with projections for hypothetical programs at the American Institute of Physics (aided by Maizell[33]), suggests the following general conclusions:

1. Prerun—that is, editorial and composition—costs for an abstract journal, at 1969 rates, are likely to be in the range \$8–12 per abstract, depending on such things as conscientiousness of coverage, depth of indexing, country in which editorial and composition work is done, and the like.
2. Ordinarily, more than half of these prerun costs are editorial, and less than half are composition.
3. Runoff costs (printing and distribution—see Section 4.2) are likely to be little more than half a cent a page per subscriber and, even for the largest circulations shown in Figure XIV.15, will not amount to more than a small fraction of the prerun costs.

For a titles journal, designed for current awareness and hence without indexes, prerun costs might be only about one third as great as those of an abstract journal with the same coverage. But if a titles journal and an abstract journal are prepared together (as are *Current Papers in Physics* and *Physics Abstracts*), most of the editorial work can be shared, and the incremental cost of the titles journal should be many times smaller, being only a little more than the composition cost.

We have been talking so far about traditional methods of printing: hot metal and letterpress, or typewriter and photo offset. (The latter, being more economical, was adopted some time ago by most of the abstract and titles journals of interest to us.) But modern techniques of computer-controlled composition—most commonly photocomposition of material for production of offset plates—make it possible for many useful products, with different arrangements or selections of the same material, to be produced with only a single composition (keyboarding) operation, though this operation may be a bit more expensive than mere typewriter composition. *Physics Abstracts* is already making use of this technology. The American Institute of Physics is currently undertaking to provide a wide range of secondary services that take advantage of this technology as well as of the type of pooling of editorial effort mentioned in the preceding paragraph. For a sizable portion of the primary literature that these services cover—the AIP-originated journals and their translations of Russian journals, which together amount to over one third of the world's physics literature—the composition cost for titles, abstracts, and the like need not be charged to the secondary service, as the material has to be composed anyway to produce the primary journal. In no other field of science is the situation so favorable for this sort of integration of primary and secondary services. Even when it is necessary to recompose a title and abstract from a foreign journal, however, a valuable economy results from being able to use the same composition for a titles journal, an abstract journal, a journal of "advance abstracts" circulated before appearance of the primary article, and, of course, a computer tape service from which any purchaser can construct his own information system.

Table XIV.6 gives details of these new secondary services planned by the American Institute of Physics[56] and shows their relation to those of the Institution of Electrical Engineers, with which AIP cooperates. In general, the division of labor is based on the following principles:

1. AIP services are aimed at *selective* coverage of the primary literature,

TABLE XIV.6 Secondary Services Offered or About To Be Offered by the American Institute of Physics and the Institution of Electrical Engineers

Service	Description[a]	Prepared at
With abstracts		
Physics Abstracts	Comprehensive abstract coverage of the world's physics literature (also available in microfiche)	IEE
Current Physics Advance Abstracts	Abstracts of articles in leading journals (initially AIP only), issued prior to publication in these journals	AIP
INSPEC Physics Tapes	Magnetic tapes with the same data as *Physics Abstracts*	IEE
Searchable Physics Information Notices (SPIN)	Abstracts, citations, and subject classifications for articles in about 70 leading journals	AIP
Without abstracts		
Current Papers in Physics	Comprehensive title listing for the world's physics literature, arranged by subject	IEE
Current Physics Titles	Three journals (*Nuclei and Particles, Atoms and Waves,* and *Solid State*) listing titles and key words for articles in about 70 leading journals, arranged by subject	AIP
Uncertain		
Current Physics Bibliographies	Specialized bibliographies in relatively narrow area, periodically updated (planned for 1973)	AIP

[a] In addition to the characteristics listed, all the AIP services supply a cartridge and frame number for location of the articles in their primary service *Current Physics Microform,* a film-cartridge form of the full texts of all papers covered in *Current Physics Advance Abstracts.*

IEE services at comprehensive coverage. Thus the AIP services will cover only articles in several score of the most important journals.

2. AIP plans to integrate the input from its own primary journals with the production of those journals and will supply tapes of the secondary information so obtained to IEE for use in IEE's services, while receiving from IEE taped information from other journals.

It is clear from the table and from what has been said above that a well-planned and efficient set of secondary information services must involve many interrelationships not only among its own elements but also with primary publications, review literature, and other components of the communication picture. (We shall return to this broader picture in Chapter 9.)

This factual description of recent and prospective developments in secondary services has already revealed a partial answer to the question posed at the start of this subsection: The societies that produce the basic secondary services need to provide a variety of such services and to minimize costs by integrating them with one another and with primary journal production and by sharing information with other secondary services. It is to be hoped that further advances in the cooperation of different services and in establishing compatibility of tape records will eliminate much of the wasteful duplication of effort that still exists. Such economies are important for two reasons: One is that the availability of these basic services to the individual scientists who use them is, as we have seen in Sections 3.3 and 3.5, significantly dependent on their price; the other is that special groups of users may have need for a wide variety of tailor-made secondary services, and one should make it as easy as possible, economically, for such "user-group" services to be put together from the material available in the basic services, for example, by commercial enterprises or by large research organizations.

Despite the economies we have just referred to, however, the prerun costs—editorial, indexing, and composition—will undoubtedly continue to be predominant in the production of abstracting and indexing services and to be large enough to price many potential buyers out of the market if they must be recouped from subscription income (see Figure XIV.15). This is an unhappy state of affairs: If the value of the product to a buyer, and through him to society, is greater than the cost of producing one *additional* copy for him (that is, the runoff cost), society as a whole gains through providing this copy to him; yet it will not be provided if the market price is greater than the value to him, since he will be willing to pay no more than its value to him. When there is a large discrepancy between the runoff cost and the market price—for a nonprofit production, the sum of prerun and runoff costs per copy—there will be many potential buyers in this range, and the loss to society will be considerable. This is the case with the large abstract journals (except *Nuclear Science Abstracts*) today.

In a situation of this sort the obvious remedy—short of a system of discriminatory pricing so ingenious that it makes each user pay approximately what the product is worth to him—is some sort of subsidy from a central source of funds. Subsidies, of course, can be of many kinds: One can subsidize via the input, as is done for primary journals with page changes. This could be done for secondary services via an "abstract charge," levied by the primary journal along with its page charge; however, it would be difficult to implement this policy for more than a minor fraction of the world's literature. At the other end, one can subsidize the purchaser of the product. The pros and cons of this type of subsidy have been discussed elsewhere in relation to primary journals (SATCOM,[4] pp. 198–204); for secondary services offered by nonprofit organizations some of the disadvantages disappear, but others remain. Finally, of course, one can subsidize the secondary service directly. This is feasible, but it entails bureaucratic dangers and probably would require international bargaining to ensure an equitable sharing of subsidy costs. Since the issues involved are not peculiar to physics, we shall not attempt to prescribe a solution here but shall merely conclude with this statement, the last part of which is based on our discussion in Section 3.5: *It is socially desirable to find some practical means for subsidizing the prerun costs of abstracting and indexing services; in physics, the net benefits of such subsidy would be quite appreciable.*

At some future time, it may become possible to provide secondary services that use in considerable detail the information contained in citations. As distinguished from more taxonomic fields of science, subject indexing in physics requires enormous expertise and would be very expensive to do satisfactorily, though it has always been done and although more ambitious attacks on it are currently being mounted. On the other hand, keyboarding of citations is a purely clerical operation, so if ways can be found to classify papers on the basis of the papers they cite, one can have the benefits of much high-quality intellectual input (that of the authors) at low cost (unless too much computer time is involved). Some very exciting progress in computer processing of citation information has been made by Schiminovich and others[57,58] at the American Institute of Physics, but many problems remain to be solved before a service for the rank and file of users, more sophisticated than that already provided by simple citation indexes, can be evolved.

As we have indicated earlier, secondary information is now available on computer tapes that can be purchased. The American Institute of Physics and the Institution of Electrical Engineers offer such tapes covering all of physics (see Table XIV.6); the Institute for Scientific Information offers tapes covering all science and technology; nuclear-physics information is available on tapes from the Atomic Energy Commission and the International Atomic Energy Agency. Major disciplines bordering physics (for example, Chemical Abstracts Service) often offer such tapes. A catalogue of available tape services has been prepared by the American Institute of Physics.[59] From any or all of these, any institution with a sufficiently large stake in physics information can put together, with purchased or home-generated software, a secondary information service geared to its own needs. Many such systems are in operation in chemistry, far fewer in physics. Particularly noteworthy is the SPIRES system at Stanford University, which now covers high-energy physics and is being extended to other areas; development is being guided by particular attention to user response.[60] Still another type of service is searching of central tapes on request, as typified by the Euratom information service based on a thesaurus indexing of abstracts (see, for example, Anthony *et al.*[38]).

In the more distant future, an interactive querying of computer files of the physics literature may become available to physicists generally. Such a capability has already been used experimentally in physics ("Project TIP"—see, for example, Kessler[61]) and is becoming operational for an engineering part of the library at MIT[62] ("Project INTREX"); further examples are listed in Chapter 7 of the SATCOM report.[2] But in the present state of technology, these systems seem too expensive to serve as a full-fledged substitute for printed abstracts, indexes, and citation indexes, and there are not enough data on user response to enable their undoubtedly real benefits to be assessed quantitatively. Improvements and further experimentation should be vigorously pursued.

4 Primary Publication

We turn now to what is really the heart of communication in any science—primary publication. As Figure XIV.3 and especially Figures XIV.4, XIV.5, and XIV.7 show, research journals and their informal counterparts, preprints and reports, play the central role in the communication of the details of ideas and data and may well be predominant even for attention focusing. In this chapter, we shall discuss these two forms of communication, and also the communication of new research results in books such as conference proceedings or monographs containing previously unpublished material.

A wide variety of facts and statistics about the primary literature are available, which we have tried to order as follows: First, in Sections 4.1 to 4.3, we shall discuss characteristics of research journals and the economics of their production. Section 4.4 will be devoted to primary publication in books. As a great deal of information is available from studies of formally published articles in journals or books, for example, from statistics from abstract journals, we shall next devote section 4.5 to a survey of some characteristics of research work that are revealed in such studies. Section 4.6 will deal with preprints and reports. Section 4.7 will discuss the problem of assessing quality in different areas of publication. Section 4.8 will consider a number of topics relating to patterns of use of published literature. Finally, in Section 4.9, we shall survey the whole picture and its future outlook.

As the subsections just mentioned are quite numerous, it is worthwhile to mention here at the outset several sources of information, each of which relates to a number of different subsections. One is the review of the literature of physics by Anthony *et al.*,[38] already referred to in Chapter 3. Another is a survey of primary journals in all fields,[4] prepared in connection with a study of their economics. A third[63] is a study of a particular journal (*Czechoslovak Journal of Physics*). Two more are studies, by the American Institute of Physics, of the literature covered by *Physics Abstracts*.[49,64] A final three are studies by the Abstracting Board of the International Council of Scientific Unions.[45,46,65]

4.1 RESEARCH JOURNALS: NUMBER, BULK, PRICE, AND CIRCULATION

It will be well to start our discussion of research journals with a quick look at their number and diversity, on a worldwide scale. We have already seen, in Figure XIV.13, how slowly the number of articles deemed relevant to physics diminishes as one scans larger and larger numbers of journals. It is illuminating to look at the list of journals ranked in the order of the number of articles they contribute to the *Physics Abstracts* listings[49]: Figure XIV.17 shows how these journals are divided between what could properly be called physics journals and journals of other types. It is clear that the number of true physics journals that contribute appreciably to the world's literature is not very large—only a little over a hundred—but that there are a huge number of journals in nonphysics fields that occasionally have articles with some physics content. Let us focus for the moment on the physics journals. Figure XIV.18a shows how 64 typical physics journals distribute over various ranges of size; Figure XIV.18b shows how these same jour-

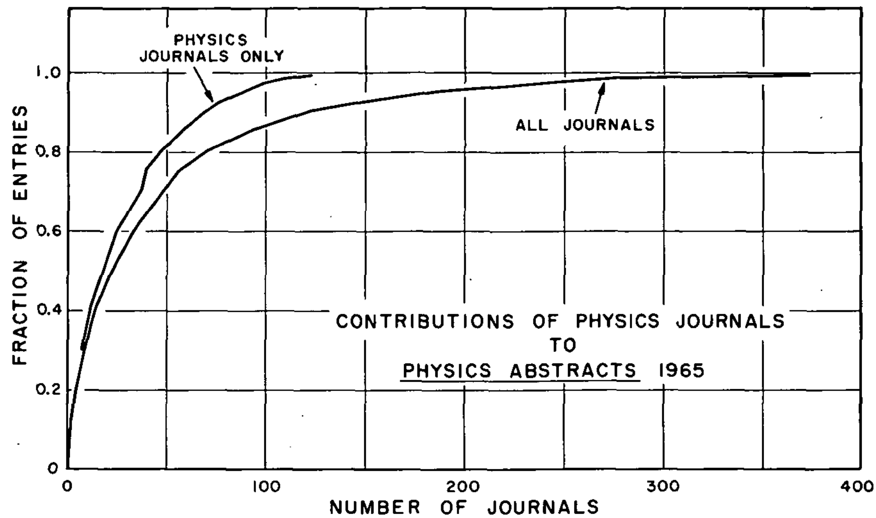


FIGURE XIV.17. Contribution of physics journals to the entries in *Physics Abstracts* 1965 compared with that of all types of journals. The lower curve is that of Figure XIV.13, its ordinate being the fraction of all entries coming from the n most prolific journals, where n is the abscissa. For each point on the lower curve, the point with the same ordinate on the upper curve gives the number n_p of those n journals that are devoted primarily to physics topics. The differences $(n-n_p)$ consist of about three fifths of journals predominantly devoted to a single nonphysics discipline and about one fifth each of multidisciplinary journals and journals devoted to narrow specialties on the periphery of physics.

nals distribute in price per unit amount of material. Although the sample is a rather sketchy one, it illustrates several points on which we can add further comments:

1. Society-run journals in physics tend to be large, especially in the United States, and tend to dominate the publication scene in physics; this situation is not true of most other areas of science. An examination of the contributions of the 125 journals with the largest contributions to *Physics Abstracts*[49]—these contributed 90 percent of all entries—shows that 67.6 percent of the entries came from physics journals, as defined in connection with Figure XIV.13, and that 53 percent of the entries from these physics journals came from society journals in non-Communist countries (36 percent were from AIP nontranslation journals alone), 25 percent from journals of Communist countries, 21 percent from commercial journals, and only 0.5 percent from journals of other types. Since a large proportion of

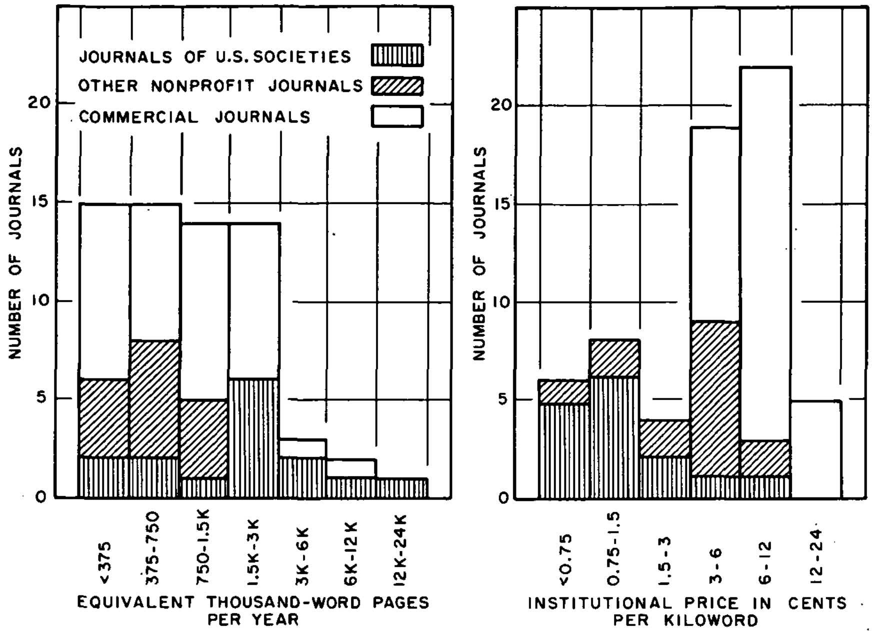


FIGURE XIV.18 *Left:* Distribution of 64 typical physics journals of non-Communist countries in annual bulk as of 1968. *Right:* Distribution of 64 typical physics journals of non-Communist countries in price per kiloword as of 1968.

the articles from Soviet sources reach Western readers through AIP translation journals, not counted in the figures just given, the role of societies as publishers looms very large indeed.

2. The prices (to institutions) of physics journals vary enormously: In 1968 the most expensive of the commercial journals cost about 21¢ per kiloword, as compared to 0.22¢ per kiloword for the *Physical Review*. (The discrepancy is less now, though still large: For example, in 1970, the *Physical Review B* sold for 0.66¢ per kiloword.) For the sample shown in Figure XIV.18, the median 1968 price for U.S. society journals (mostly with page-charge support but including some translation journals) was below 1.5¢ per kiloword, that for foreign (non-Communist) society or subsidized journals (only a few with page charges) was about 3¢, and that for commercial journals about 8¢.

Prices are of interest for two related reasons, both having to do with their relation to circulation. As we have noted briefly in Section 2.1, and as we shall discuss further in Section 4.7, the utility of journals depends on their ready availability to those who may profit by consulting them, and this availability, as we shall see presently, is usually correlated with price.

Thus, if we are interested in making journals as useful as possible, we should look for ways to keep prices down. In addition, if we can learn enough about the relation of price to circulation, we will have a measure of the value the scientific community places on journals. For these reasons, therefore, it is appropriate to look at the circulations of typical physics journals, and their correlation with price. Figure XIV.19 shows some 1968 data for a number of U.S. and foreign journals. These data provide a par-

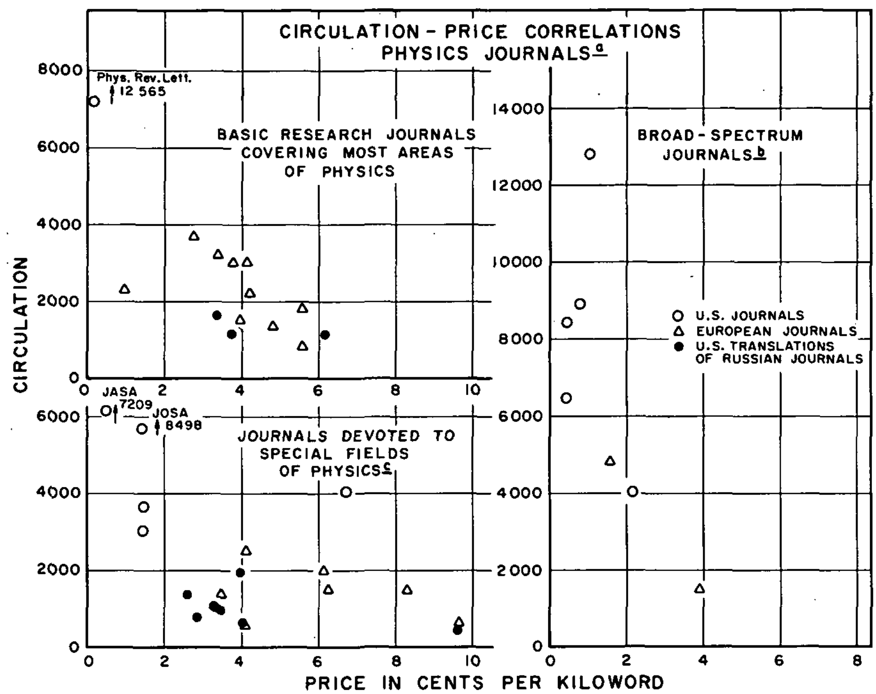


FIGURE XIV.19 Circulations of a number of U.S. and Western European physics journals of different types correlated with price (in cents per kiloword) to institutional subscribers.

[a] Although it would be better to show institutional and personal subscriptions separately, the separated data were not available for many of the journals. Data on AIP journals suggest that in the 6000–9000 range the subscriptions are divided about half and half between institutional and personal, and that below 2000 they are almost entirely institutional. Availability of data also limited the representation of commercial journals. Although many are shown on the plots, circulation figures for many others, including most of the more expensive ones, could not be obtained.

[b] Examples: *Journal of Applied Physics, Journal of Chemical Physics, Review of Scientific Instruments.*

[c] Examples: *Physics of Fluids, Applied Optics, Nuclear Physics, Journal of Mathematical Physics.*

tial illustration of the following statements, whose basis has been more fully discussed elsewhere[4]:

3. Journals of higher price tend to have significantly lower circulations. Although the correlation in the lower range of prices may be due merely to the fact that a more popular journal does not need as high a price to avoid going in the red, the correlation at the higher prices undoubtedly represents buyer resistance; this resistance comes partly from individual subscribers, partly from the cutting out of multiple institutional subscriptions, and partly from institutions whose interest in the field of a journal is marginal. It is worth noting that of 16 journals with physics material, identified by Cooper and Thayer[66] as being held by more than 250 libraries listed in *Access*, all but one were priced near or below 2¢ per kiloword.

4. Despite buyer resistance, enough libraries will purchase physics journals at quite high prices (up at least to the 21¢ per kiloword quoted above) to protect any journal with reasonable content from the likelihood of failure as a result of lack of subscribers.

5. Journals devoted to areas of physics that border on chemistry, engineering, and like disciplines (see Figure XIV.19, parts *b* and *c*) tend to have especially large circulations. In similar accord with one's natural expectation, journals devoted to more narrowly specialized topics (most of those in Figure XIV.19b) have lower circulations, which vary considerably from one specialty to another.

The number of journals has been growing slowly. A crude measure of this growth is provided by a comparison of *Physics Abstracts* entries in 1934 and 1965. In the latter year, according to Figure XIV.17, about two dozen journals accounted for half the entries; in 1934, about one dozen accounted for half. A less meaningful statistic, which nevertheless has a similar significance, is the comparison of the numbers of journals scanned for the two years—about 260 in 1934 and 495 in 1965. Such figures suggest an average increase of the order of 2 percent a year in the number of physics journals, over a period including World War II. It is, therefore, not unlikely that since the war the number of physics journals has increased at a rate similar to the 3 percent or so a year that has been estimated for scientific journals generally.[4] This latter increase is an average of about 8 percent a year for commercial journals (which in physics are predominantly European) and 1.4 percent a year for nonprofit journals.

The growth in numbers of journals just mentioned is much slower than the growth in volume of publication typified by Figure XIV.10, which is about 11 percent a year, or than this figure correlated for expanding coverage of the abstract journals (at least 7 percent a year). Thus the physics journals must, on the average, have been getting fatter. This is well known to be the case for the larger journals; Figure XIV.20 illustrates the growth over the last several decades for a number of physics journals. Also drawn on the figure is a line showing the growth of U.S. PhD physics manpower

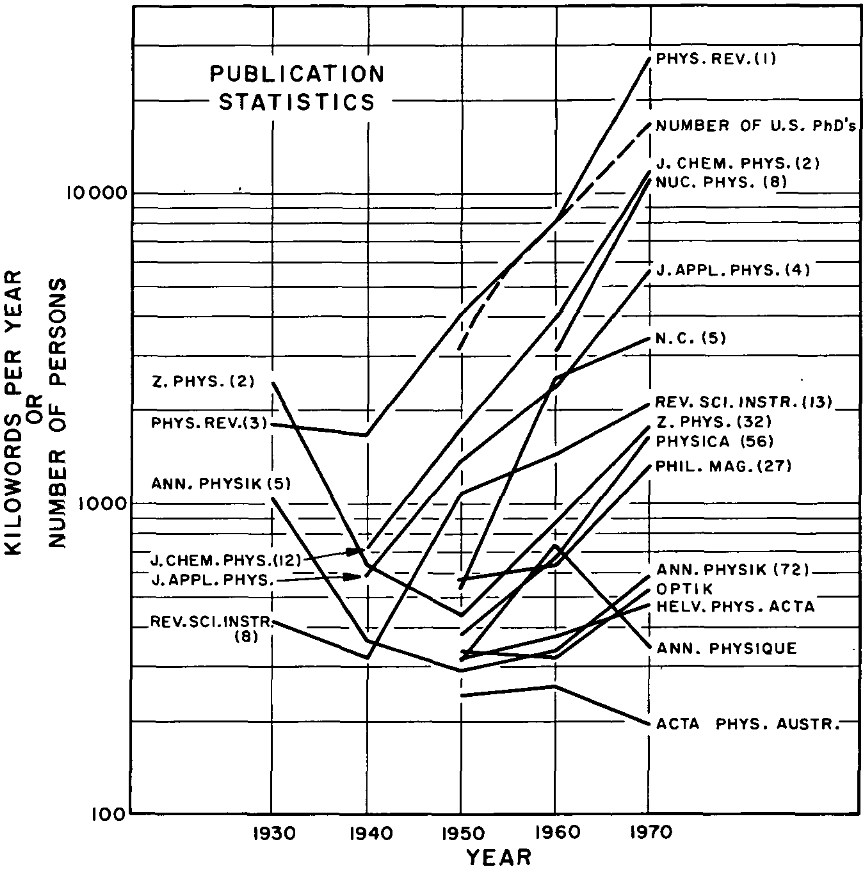


FIGURE XIV.20 Changes from decade to decade in the amounts of material published by various physics journals. The numeral appended to each journal at the left is its rank order in number of articles abstracted by *Physics Abstracts* in 1934; the numeral at the right is the corresponding rank order for 1965. If we exclude interdisciplinary journals (*Comptes Rendus, Nature,* and the like), letter journals, nonphysics journals (for example, *Journal of the American Chemical Society*), and Soviet journals, the journals shown include all others in the first ten of the 1934 ranking, or in the first 13 of the 1965 ranking. Some smaller journals are also shown for comparison. The dashed curve gives the growth of the number of physics PhD's in the United States.

over this period; we shall comment presently on the comparison of the slopes of the dashed and solid lines.

Despite the increase in bulk of journals and a slight increase in their price per page, the circulations of most scientific journals published by societies rose, rather than fell, in the 1960's.[4] Data on commercial journals are not available. These might show a different picture, as their prices are higher, but one can at least say that so far, in the field of physics, none of them have failed financially, nor have their financial prospects worsened enough to discourage the initiation of a number of new journals in the last few years. Although physics journals have not done as well in this respect as those of some other fields—the *Physical Review*, for example, has declined slightly in circulation—it is noteworthy that even in the interval 1968–1970, when there was a universal shortage of funds, about as many journals of the American Institute of Physics rose in circulation as declined.

It is sometimes said that the rate at which physicists publish has been increasing in recent years, that is, that the average physicist publishes more papers per year than he used to. Such statements have usually been based on the fact that the rate of growth of the number of entries in *Physics Abstracts*—in the last decade about 12 percent a year, or a doubling time of six years—is considerably greater than any reasonable estimate of the growth of the world's physics manpower. (In the United States, the doubling time for total physics PhD's in the decade was a little under nine years—see Chapter 12.) But this comparison needs correction for the fact that over this period the coverage of *Physics Abstracts* has been extended to conference proceedings, some reports and theses, and other types of media and has broadened in the fringe areas of physics. An even less reliable indicator, sometimes used, is a comparison of pages or papers in the *Physical Review* with the membership of the American Physical Society, or, as in Figure XIV.20, with U.S. PhD manpower. We have tried to get a more reliable measure of productivity by taking random samples of the membership of the American Physical Society in 1955 and 1970, noting how many papers by these authors in journals were listed in *Physics Abstracts* for the same year, and noting how many of these were in fringe areas in which the coverage of *Physics Abstracts* might have changed. The result was striking: The average society member was an author of 0.28 paper in journals covered by *Physics Abstracts* in 1955 and of 0.54 in these and similar new journals in 1970. A part of this increase has, of course, been due to the rise in multiple authorship of papers (see Figure XIV.34) rather than to a greater total production per member. But even when we correct for this factor by taking averages of about 1.8 authors per paper in 1955 and 2.3 in 1970, we still are left with about 0.15 paper per member per year in

1955, and 0.23 in 1970. The cause of this increase is not yet clear; one may speculate whether it is related to an increasing concentration of physicists in academic institutions and the involvement of an increasing fraction of these institutions in research.

The alert reader may have been puzzled by the fact that the product of 1970 papers per APS member (0.23) multiplied by number of members (28.8 thousand) is considerably less than the number of 1969 published research papers from U.S. institutions (13,500; see Tables XIV.10 and XIV.11), the great majority (85–90 percent) of which must surely have been in journals qualifying for inclusion under the criteria used by *Physics Abstracts* in 1955. One major explanation of this finding is that quite a large proportion of the authors on physics papers from U.S. institutions (slightly over half the authors, according to counts in a limited sample) are not members of the American Physical Society. Another comparison that requires explanation is that the publication rate reported by Hagstrom[31] for a sample of physicists was over three times higher than the rate reported here. In this case, the explanation has to do with a difference in the populations; ours was taken from all APS members, his from the staffs of PhD-granting universities.

But to what is the increase in output per physicist due? Further studies are needed to sort out the roles of several possible factors. Have physicists become more efficient at doing research because of improved tools or funding? Has the rise of letter journals and published conference proceedings encouraged extensive duplicate publication? Has the proliferation of the literature led to an analogue of the "cocktail party effect"[67] in which everyone has to shout to be heard above the background and, in turn, raises the background for others? One possible explanation is that the fraction of physicists in academic institutions—where incentives for publication are greater—has increased. Further, the percentage of institutions of higher education that are engaged in research has increased. Consequently, more physicists are employed in research-oriented institutions.

4.2 RESEARCH JOURNALS: PRODUCTION AND ECONOMICS

Although, as we have just noted, it always seems possible to move the finances of a physics journal from the red to the black by raising the subscription price, the economics of journals have often caused great anguish to publishers, buyers, and others. And rightly so: The $21 million or so that is now spent each year in the United States for primary physics journals plus associated library services (see Section 2.3 and Figure XIV.9) strains the budgets of many libraries and departments; and the scientific societies that publish most of the U.S. journals have been properly loath to let the availability of their journals be lessened by the loss of individual and institutional subscriptions that a rise of price would entail. While the same forces have been acting in nearly all fields of science, the issues have emerged more clearly in physics than in any other field, because of the unusually large share of the journal production work that is in the hands of the American Institute of Physics and its Member Societies. In a field in which the picture is the reverse—publication dominated by a diversity of commercial journals—the societies never think about the economic problems of the journals in their role as authors, since there are no page charges; they often do not regard their libraries' financial problems as their own, and if they find that they have to go to a main library to find a journal they formerly could peruse in the next room, they are apt to consider the causes of this change as too remote from themselves to be worth investigating. But in physics, prices, page charges, and other policies are set by societies controlled by the physicists themselves, and they must face the issues.

What determines the cost of producing a journal? Table XIV.7 shows the various operations that are involved, arranged in rows according to the location or organization in which they are performed and in three columns according to the nature of their dependence on the bulk, circulation, and other characteristics of the journal. The left-hand column, labeled "prerun," contains all operations that are necessary before production of the first copy of the printed research or development work that we presume to be the principal content of the journal and its raison d'être. These costs are independent of the number of copies to be produced, but for a given type of material, they increase with the number of research pages (or of papers) published, and for a sufficiently large operation they are proportional to this number. The right-hand column, on the other hand—the one labeled "runoff"—contains operations whose costs depend on the number of subscribers to the journal, being proportional to this number if it is reasonably large. Most of the runoff costs, of course, also increase almost proportionally with the page bulk of the journal; however, subscription maintenance—keeping records of subscribers, billing them, and the like—does not, while covers, wrapping, and mailing may do so only in a stepwise manner. The middle column of the table contains operations that can be viewed as incidental rather than as necessary to the publication of new knowledge and that need be performed only insofar as they are regarded as independently desirable or can be made to yield more income than they cost. For U.S. research journals in physics (as contrasted with the news-oriented journal, *Physics Today*, or with many research journals in medicine), advertising and the other activities of this middle column play a minor economic role.

It would take too long to discuss each component of each column of Table XIV.7; the interested reader can find details in the SATCOM Task

TABLE XIV.7 Elements of Cost in the Production of a Journal

Location	Prerun Costs	Miscellaneous Costs	Runoff Costs
Editor's office[a,b]	Technical editing[c] Salaries of editors Clerical staff Telephone, postage, etc. Referees		
Production Office[a,b]	Copy editing[d] Copy editors' salaries Clerical staff Art Department Indexes Miscellaneous Contribution to proof-reading Page-charge billing [Typewriter composition[e]]	Promotion Advertising Solicitation and correspondence Processing of copy News, etc. Handling of reprint and back-number orders	Subscription maintenance
Printer	Composition[e] Typesetting Proofreading	Back Numbers Overrun for back-number stock Mailing Reprints	Wrapping and mailing Printing, etc.: Paper Presswork Binding
Engraver	Engravings		

[a] Editor's office and production office may sometimes be operated together.
[b] Note that overhead and employee benefits should be included.
[c] Technical editing includes receipt of manuscripts and all work involved with decisions as to their acceptability, need of revision, etc.
[d] Copy editing, the preparation of manuscripts by the typesetter or other compositor, includes such things as marking them for the compositor, standardizing headings and footnote arrangements, planning the layout of figures and tables, etc.
[e] Typesetting may sometimes be replaced by typewriter composition done in the production office, plus plate preparation done by the printer. Computer photocomposition, again outside the printing house, may also be used.

Group report,[4] from which the data to follow have been taken and, for journals of the American Institute of Physics, in a recent report by Koch.[68] Here it will suffice to sketch the overall picture and call attention to a few facets of it. Consider first how the total cost of producing a journal is divided between prerun and runoff costs. Typical values of these costs for journals in a variety of fields of science are shown in Figures XIV.21 and XIV.22, referred in each case to a kiloword of material. The variation in prerun costs from one journal to another is in part an artifact of bookkeeping—accounts are usually kept by rows rather than columns of Table XIV.7, so that determining the true prerun costs may be difficult; in part, it is a manifestation of hidden subsidies—editors' time, office space, and the like

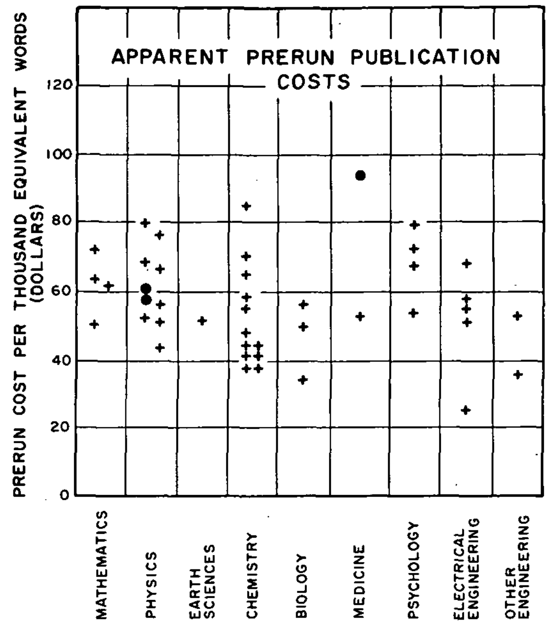


FIGURE XIV.21 Distribution of apparent prerun costs (as defined in Table XIV.7) per unit amount of material published, as they appeared on the books of a sample of U.S. journals, in various fields. For journals containing advertising or other non-research material, the figures are *intended* to apply to the material only, but may not always fully realize this intention. Black dots identify journals with full-time paid editors, for which one can be sure that there are no hidden costs for technical editing.

that are supplied gratis in some cases; and in part, it is a reflection of real differences in such things as amount of editorial effort expended or composition costs. Physics articles, with their high content of mathematics, are

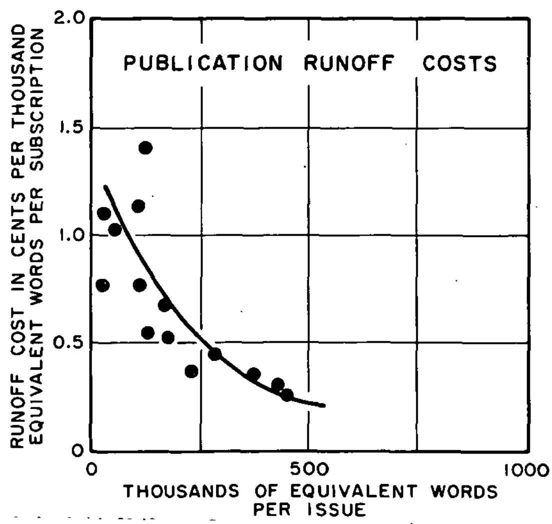


FIGURE XIV.22 Runoff cost (as defined in Table XIV.7) per unit amount of material printed versus average size of a single issue, for a sample of journals in all fields. The solid curve is an arbitrary one drawn for reference. [Source: SATCOM.[4]]

particularly expensive to set in type; on the other hand, many (by now a majority) of the AIP journals have preceded those of most other fields of science in adopting typewriter composition, which is less expensive than set type. (Most of the data of Figure XIV.21 antedate this change, however.)

Using Figures XIV.21 and XIV.22, let us see how the prerun and runoff components of the production cost look in typical cases. If $n(p)$ is the number of subscribers at price p,

$$\text{production cost} = s + rn(p), \tag{4.1}$$

where s is the total prerun cost and r the runoff cost per copy. As we have seen in Figure XIV.19, the more widely circulated of the U.S. journals in physics have circulations of the order of 6000–12,000. With typical values $r \approx 0.3$–0.4¢ per kiloword per copy, the runoff term of Eq. (4.1) then ranges from \$18 to \$48 per kiloword, that is, from a minor fraction of the typical prerun cost of \$60 per kiloword to a major portion of it. For most of the commercial and foreign journals, with their smaller circulations, $rn \ll s$. For journals that try to meet all their costs from subscription income, this dominance of prerun costs has two unfortunate consequences:

1. It necessitates setting a price p much higher than the runoff cost r. This is a deprivation to the class of buyers to whom the value of the journal exceeds the cost r of producing an extra copy but is less than the prorated cost $r + (s/n)$: Society as a whole would gain if extra copies were run off and provided to these buyers, yet the set price excludes them from the market.

2. The finances of the journal become unstable. If the amount of material it needs to publish suddenly increases, costs may rise alarmingly before new subscription rates can be set. Alternatively, if the number of subscriptions fluctuates downward, the planned balance of costs and income can be drastically upset.

These troubles are, of course, not peculiar to physics; they beset journals in all fields. The physics community can take credit, however, for introducing, in 1930,[69] a method of dealing with them that has proven so satisfactory that it has come to be adopted by the majority of U.S. society journals in all the other sciences. This is the so-called page-charge practice, whereby the institution sponsoring a piece of research pays the journal for its publication at a rate that now generally covers most or all of the prerun costs. The charge is not compulsory, and payment is not expected if it would constitute a hardship. The decision to publish a paper is made by the editor without knowledge of whether the page charge will be paid. The average subscription price is maintained at a level that exceeds the runoff cost by enough to support the prerun costs of those papers that do not honor the page charge.

However, for two major reasons, the great majority of papers in journals of the American Institute of Physics honor the request for page-charge payment: First, a policy enunciated by the Federal Council for Science and Technology[70] allows such payments to nonprofit journals to be charged to research budgets for all federal grants and contracts, and program officers of the major agencies have encouraged applicants to include budget items fo publications; second, the AIP staff has worked diligently over many years to persuade the management of industrial and other research organizations that these organizations have a civic duty to contribute their share to the support of publication. Since over half of the articles published in U.S. physics journals have federal government support, and nearly half the remainder are of industrial origin—foreign and miscellaneous U.S. papers make up the other half—there should be available funds for page charges to U.S. journals in most cases. Foreign journals only rarely use page charges; because of the fragmentation of their material over sources of many nationalities, it would be difficult to get adequate governmental and other policy commitments to ensure a high level of honoring. Moreover, most European journals are issued by commercial publishers.

It is not difficult to see that page charges can provide a very effective answer to the problems enumerated as 1 and 2. That this practice has made low prices and high circulations possible is attested by Figure XIV.19, where the circles represent U.S. journals with page charges and most of the triangles represent journals without page charges. The contribution to stability results from the fact that a fluctuation in the number of papers submitted is accompanied by a corresponding change in the page-charge income, which will take care of most of the change in prerun expense; similarly, when there is a fluctuation in the number of subscribers, the changes in subscription income and in runoff expense will be comparable. In either case, the finances of the journal remain roughly in balance. Therefore, the journal publishers can afford to take a cooperative attitude toward Xeroxing, reprinting, and other uses of their product that may be helpful to the scientific community.

Figure XIV.23 shows the experience of journals of the American Institute of Physics over the last decade in regard to the honoring of page charges (Koch,[68] updated). The downturn occasioned by the scarcity of funds in 1968 is very evident, but it has been effectively arrested following introduction, by most of the journals, of the "two-track system"—a policy of publishing nonpaying papers only as rapidly as the journal's budget permits, so that if too many such papers occur, they will face a delay in publication. For most of the journals that used the two-track system, the delays have

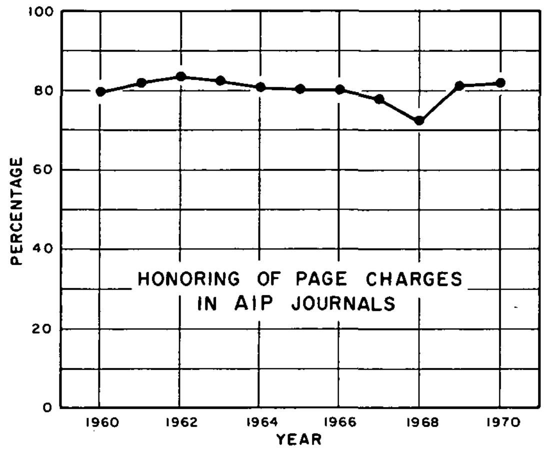


FIGURE XIV.23 Time changes in the percentage of papers published in AIP journals for which page charges were honored.

been less than three to four months; in one case, however (*Journal of Chemical Physics*), there was at one time a delay of 11 to 12 months for nonhonoring papers.

Despite its general success, from time to time some people have objected to the page-charge system. The issues raised have been summarized elsewhere (SATCOM,[4] pp. 10–14); we shall comment here on only two of them. One is the argument that they impose a hardship on impecunious institutions, especially academic ones. We feel that this view misunderstands, on the one hand, the freedom of those without funds to forgo payment and, on the other hand, the fact that a nationwide shift toward publication in the often much more expensive non-page-charge journals could in some cases be more expensive (as well as providing poorer dissemination) for U.S. colleges and universities, since the drop in physics-department expenses would be compensated by a rise in library outlays. We have indicated above why marketing of journals at near runoff cost gives society as a whole the best return for its money, and we feel that a flexible and commonsense administration of the page-charge system is the best way of making this possible. (For further details, see SATCOM,[4] Appendix Section IV.)

The other argument on which we wish to comment is the contention that journals lose their incentive for efficient operation when so much of their income is assured through page charges. Although this argument sounds superficially plausible, a detailed look at the economic pressures on a producer of journals (SATCOM,[4] Appendix Section IV A) shows that page-charge subsidy does not adversely affect most of the pressures toward efficiency. Even more convincing is a comparison of production costs of journals with and without page charges. For example, Figure XIV.24 compares the total cost (prerun plus runoff) per kiloword per subscriber for a number of non-page-charge physics journals with that of the *Physical Review* and with a curve showing what the *Physical Review* would cost if it printed an arbitrary number of copies at the same runoff rate (from Eq. 4.1). As many of the comparison journals are nonprofit, are produced in countries where labor costs are lower than in the United States, and sometimes have hidden subsidies in the form of editors' time and space, it is obvious that the *Physical Review* cannot be accused of inefficient operation.

A good deal of thought has been devoted to computer-controlled composition. Use of this mode for titles, abstracts, and references seems feasible, and, as we noted in Section 3.6, integration of the composition of primary and secondary journals is likely in the next few years as far as this material is concerned. But for full text, where there are no such opportunities for economizing via multiple use, computer composition is not yet competitive with simple typewriter composition for offset; the handling of mathematical expressions poses special difficulties, though some progress has been made toward overcoming them. The intriguing prospect of a major saving in composition costs by use (with editing) of computer tapes prepared by authors in their own institutions—that is, by using essentially one keyboarding for author's manuscript, preprints, and journal composition—must be relegated to the remote future.

(We shall return to some broader economic issues in Section 4.9.)

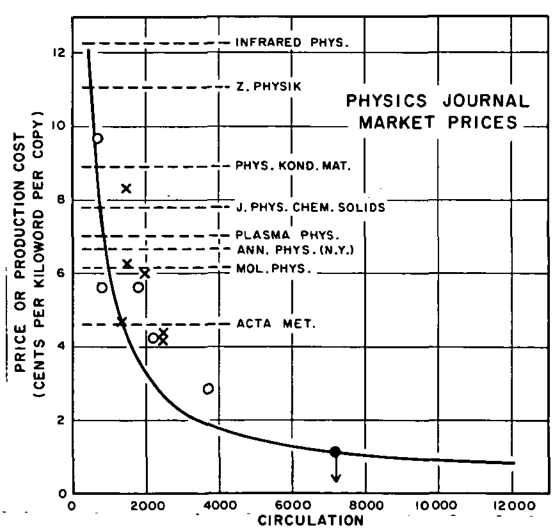


FIGURE XIV.24 Comparison of market prices of various physics journals in 1968 with the curve of what the production cost of the *Physical Review* would have been at any given circulation. The black circle is the total production cost of the *Physical Review* at its actual circulation; the head of the arrow beneath it represents the price charged to institutional purchasers; open circles are price–circulation points for various European nonprofit physics journals (without page charges); the crosses are similar points for various European commercial physics journals. The dashed lines are at heights representing the prices of a few other commercial journals for which circulations are not available. [Source: SATCOM,[4] p. 196.]

4.3 RESEARCH JOURNALS: MISCELLANEOUS CHARACTERISTICS

In addition to the topics already discussed, there is a diverse assortment of further noteworthy facts about the research literature and the media in which it is published. We shall consider here that part of the assortment that has to do with primary journals as such and the grouping of articles in them, leaving to Section 4.5 the discussion of characteristics of the research papers considered one at a time, irrespective of the medium of publication, and to Section 4.8 the discussion of patterns of use. The subjects of the following paragraphs have no logical order: We shall arbitrarily start with mechanical items (typography and format, time lags), continue with language distribution and referencing, and conclude with some aspects of intellectual structure (specialization of journals, citation patterns).

Typography, format, and the like are important for two reasons: Different choices among them can entail different production costs and can affect the efficiency with which the readers of journals use them. One variable is page size. Nearly all the journals of the American Institute of Physics, like those of the majority of U.S. scientific societies in other fields,[4] use a fairly large page, with about 1050 equivalent words per page (that is, a page set in continuous textual material would contain about 1050 words). Most foreign physics journals, and in particular the commercial ones, use a smaller page, 500 to 600 equivalent words, on the average. The large page is advantageous for several reasons: The fraction of waste space at the margins and around mathematical expressions is less than with a smaller page; the reader does not so often need to flip pages to refer to a previous equation or figure; there is less waste in Xeroxing.

There are also systematic differences between AIP and commercial journals in regard to size of type, the former using mostly 11 point, the latter evenly distributed between 11 and 12. Again, the smaller type is more economical and makes for less page flipping in reading mathematical or diagrammatic material. Studies of reader preferences and reading speed have shown that subjective preferences peak at about 11 point, reading speed at about 10 point.[71] Similar studies, incidentally, have shown that reading speed and efficiency are unaffected by whether or not right-hand margins are justified, that is, that the ragged appearance of typewriter-composed material (now being adopted by most AIP journals) does not entail any practical disadvantages.

Of more concern to most scientists is the time lag in publication. For journals publishing full-length papers, the current mechanics of the production process impose a minimum time delay from receipt of a manuscript by the editor to receipt of the published paper by the reader of about three months for monotype–letterpress and rather less for typewriter–offset.[4,68] Letter journals, by using typewriter composition, eliminating some of the proofreading steps, and sacrificing some economy for speed, can bring this minimum lag down to about one month. But for both types of journals, the lag is usually considerably greater because of

1. The time taken in refereeing and in the intellectual aspects of editing
2. The additional delay of a month or so entailed by having the composition work done overseas, if this is necessary for economy or for other reasons
3. Accumulation of backlogs if the rate of arrival of papers exceeds the rate at which available facilities or finances can cope with them

Median lags for some typical physics journals are shown in Table XIV.8. Physicists are fortunate in that these lags are, on the average, shorter than for journals in some other fields of science (SATCOM,[4] p. 142). Of course, the lags vary erratically, not only from journal to journal, but from time to time; for example, in 1970 the *Physical Review* was quite far behind because of staff reductions at the American Institute of Physics due to the influenza epidemic and other causes. The variation from paper to paper about the median is also large, usually because of factor 1 above (editing and refereeing) but sometimes, as in the case of journals with a two-track system, because of factor 3; Figure XIV.25 shows some typical distributions. In connection with factor 3, it is worth noting the recommendation of the SATCOM Task Group,[4] derived on the basis of rough estimates of the loss to society due to publication delays: *In use of a two-track system, the median additional delay incurred by papers for which page charges are waived should, if at all possible, be maintained at no longer than a couple of months or so.*

TABLE XIV.8 Median Time Lags[a] for Typical Physics Journals

Journal	Median Lag, Months as of ⟶	(Epoch)
Phys. Rev. Lett.	2.6	Oct. 1971
Appl. Phys. Lett.	3.2	Sept.–Oct. 1971
Phys. Rev. A	5.8	Oct. 1971
J. Chem. Phys.	6.8	Oct. 1971
Ann. Phys. (N.Y.)	12.0	July–Sept. 1971
Nucl. Phys. A	4.9	Sept.–Oct. 1971
J. Phys. E	5.4	July–Sept. 1971

[a] Interval from receipt of paper by editor to receipt of published version by a subscriber in the United States. If manuscript was revised, date of receipt of revised manuscript is used. For non-letter journals, only full-length papers are counted.

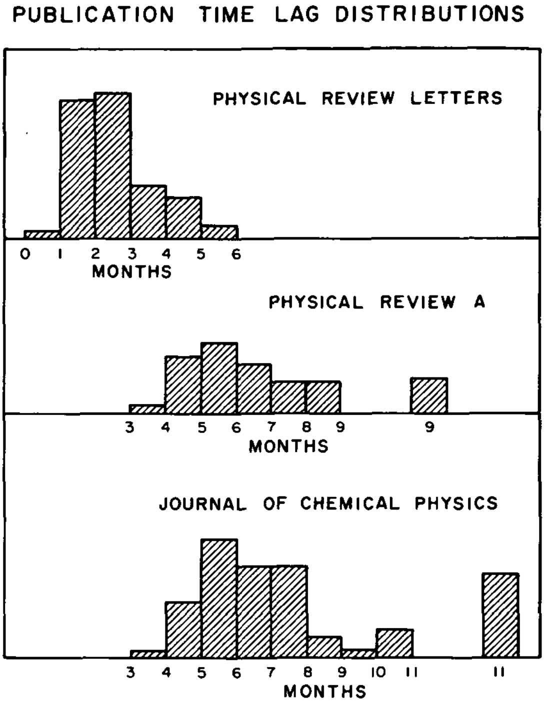


FIGURE XIV.25 Distribution of time lags in publication for typical journals (same epochs as in Table XIV.8). Histograms are not normalized.

Although the principal culprit in longer-than-minimum time delays is the practice of submitting papers to outside referees to obtain judgments on their suitability for publication, few physicists would suggest that this practice be abolished. It is used today by practically all journals. The system is sometimes criticized, however, as tending to entrench orthodoxy and to discriminate against authors who are not well known. A recent sociological study of the refereeing process[72,73] has shown that, at least in the *Physical Review* (the journal studied in greatest detail), refereeing seems to be conducted in an impartial manner. The distribution of papers over the various possible outcomes of the refereeing process is shown in Figure XIV.26. Although a journal like the *Physical Review* rejects no more than one fifth of the papers submitted to it, it is fairly certain that the knowledge that their papers will have to undergo refereeing stimulates authors to sharpen their logic, justify premises, and improve their exposition. Seen in this perspective, the custom of refereeing is, as Ziman[3] has eloquently pointed out, a key element in achieving the consensus that gives science its characteristic solidarity and its cumulative progress.

Referees usually donate their services as a voluntary contribution to the

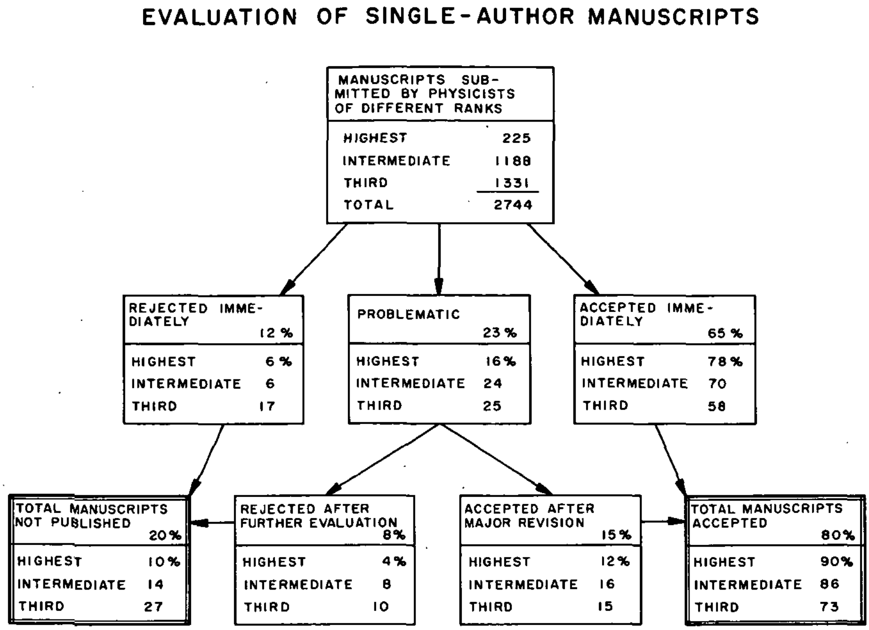


FIGURE XIV.26 Fate of manuscripts submitted by single authors to the *Physical Review*, 1948–1956, broken down according to the estimated level of prestige ("rank") of the authors. [Source: Zuckerman and Merton.[73]]

scientific community, although some commercial journals acknowledge this service with a small honorarium. The value of the time so donated must be reckoned as a hidden element in the cost of producing journals; its contribution to the prerun cost has been estimated[4] to be of the same order as the cost of all the rest of the scientific editing; like the latter, it will vary widely from case to case, values in the range of $10 per kiloword published being typical for journals with full-length papers, with rather larger values for letter journals.

Let us turn now to the distribution of journals over countries and languages and the extent to which work performed in one country is published in another. Detailed analyses of the former topic have been given,[45,46,66] from which we cite the following highlights:

1. English predominates, accounting for 68 percent of the entries in *Physics Abstracts*, probably a slightly lower percentage of those in *Bulletin Signaletique* (difficult to estimate, because in over half the cases, the language was not stated). For journals published in English-speaking countries

and in India, 98–100 percent of the articles are in English; in The Netherlands and Scandinavia, over 90 percent; in Italy and Poland, 70–85 percent; in Japan, 95 percent (*Phys. Abstr.*) or 37 percent (*Bull. Sig.*); in most of the rest of Europe, Eastern as well as Western, about half.

2. Russian is the second most used language, accounting for 17 percent of the entries in *Physics Abstracts,* rather more in *Bulletin Signaletique.* It accounts for about 95 percent of the articles published in the Soviet Union, but for only 5 percent to 25 percent of those published in the various Eastern European countries.

3. French and German each account for 5–7 percent of the entries in both abstract journals. French is used for well over 97 percent of the articles published in France and is used extensively in Belgium and Switzerland but very little elsewhere (maximum, 10 percent in Romania). German is used for 80 percent or more of the articles published in Germany but for only a minority of those published in Austria and Switzerland; it has a moderate usage in Eastern Europe.

4. All other languages account for less than 3 percent of the entries in *Physics Abstracts,* perhaps slightly more of those in *Bulletin Signaletique.* Ukrainian and Japanese seem to have the largest representations, though little material in Japanese is covered by *Physics Abstracts.*

5. Of all the journals contributing entries to *Physics Abstracts* in 1965, 37 percent accept papers in more than one language, and 80 percent publish at least some papers in English. Gross statistics like these, of course, need some qualification. For example, we have been assured by Japanese physicists that essentially all significant Japanese work in core areas of physics gets published in English, despite frequent duplicate publication in Japanese. Also, time changes in the patterns are continuing: For example, the leading German physics journals now publish rather more material in English than they did in 1964; similarly, the *Ukrainian Journal of Physics* has switched from Ukrainian to Russian.

Correlated with the language patterns of journals, but by no means synonymous, is the tendency toward segregation of publications originating in a certain nation or region into particular journals. Accompanying this trend is a tendency for physicists of one region to pay more attention to journals of the same region than to those originating elsewhere. These tendencies, which have always been present, can, at their worst, retard the progress of science appreciably. Fortunately, there are several indications that this sort of provincialism is decreasing. First, there are the linguistic indicators just discussed, especially points 1 and 5 and the discussion of point 5. In addition, there is the formation of the European Physical Society and its effort to integrate the journals of Western Europe into a set of "Europhysics Journals." The birth of highly specialized journals, usually under commercial sponsorship, has lured authors from all parts of the world to publish in the same journal. A similar effect has occurred with letters journals, because of the speedy publication they provide, and with the East German journal *Physica Status Solidi* for the same reason. Table XIV.9 gives a summary of correlations we have found in a sample of research papers from *Physics Abstracts* in 1969. These illustrate some of the points we have just mentioned and also show the influence of language and geographical distances on publication patterns.

A related characteristic of journals, on which some data are available,[45,46] is the degree of concentration of the journal literature published in any given country into a small number of journals. Figure XIV.27 shows some examples. The not unexpected trend is for countries in which a great deal of physics is published to have this publication less concentrated into a few journals.

Let us consider now the intellectual specialization of physics journals. As in most other fields of science, the relentlessly increasing volume of publication has made journals that undertake to cover all of physics grow, in many cases to such a size that individual subscriptions are both prohibitively expensive and too bulky to store. The result has been a rising birth rate for more specialized journals and a subdivision of major older journals. Figure XIV.28 illustrates the trend. Not one of the dozen journals contributing half the papers to *Physics Abstracts* in 1934 was specialized to a subfield within physics, but, by 1971, subfield journals, including subdivisions of previously general journals, have become predominant; moreover, small commercial journals seem to be viable even when devoted to what would once have been considered narrow specialties such as molecular crystals, infrared physics, thin solid films, crystal growth, molecular spectroscopy, or computational physics.

Appendix XIV.A gives a concise snapshot of the present extent of specialization of journals by listing the ten or so most productive journals in each of the sixteen subject subfields into which *Physics Abstracts* was divided in 1964. Few journals occur on many of the lists: Only one, *Comptes Rendus,* occurs on more than half of them, and only nine others occur on as many as four of the sixteen lists. (These data are from ICSU-AB[46]; similar data have been provided by Keenan and Brickwedde.[49])

One views these trends with mixed emotions. On the one hand, both common sense and circulation statistics show that neither individuals nor small research groups can continue to subscribe to, and store, vast superjournals for their private or subdepartment libraries. Moreover, as we have noted in Section 2.1, the effectiveness of communication via journals that are available in such places is unquestionably greater than that via journals that are available only in a more distant library. On the other hand, there

TABLE XIV.9 Regions of Publication of Physics Research Papers Originating in Various Regions, Determined by a Sampling of *Physics Abstracts*, 1969[a]

Region in Which Research Conducted[b]	Region of Publication: U.S.A.	U.K.	Cont. W. Europe	U.S.S.R.	Other E. Europe[c]	Africa	Japan	Other Asia	Australia & N.Z.	Other[d]	Percent Published Outside Region of Origin
U.S.A.	1116	106	227	0	16	0	4	1	0	19	25%
U.K.	57	227	69	0	5	0	1	0	0	6	38
Cont. W. Europe	110	48	618	0	32	0	0	2	0	15	25
U.S.S.R.	8	12	24	778	18	0	1	0	0	4	10
Other E. Europe	13	10	33	0	112	0	0	5	0	3	37
Africa	3	2	3	0	0	8	0	0	0	1	53
Japan	24	5	10	0	3	0	204	0	0	0	17
Other Asia	37	24	38	0	5	0	8	82	0	6	59
Australia and New Zealand	11	11	24	0	4	0	0	2	7	2	88
Percent submitted from Outside Region of Publication	19%	49%	41%	0	43%	0	6%	11%	0	–	–

[a] Criteria for inclusion in the sample are described in Table XIV.10.

[b] Location of institution where work was performed, or of first-named institution when there were several. Cases in which location of institution was not mentioned or was not easily inferrable have been omitted (these constituted about 7 percent of the papers sampled).

[c] Most of the off-diagonal entries in this column came from a single journal, *Physica Status Solidi*, published in East Germany, which has attracted many authors by its rapid publication schedule.

[d] Includes occasional journals labeled as "international" in *Physics Abstracts*, though usually published in Western Europe.

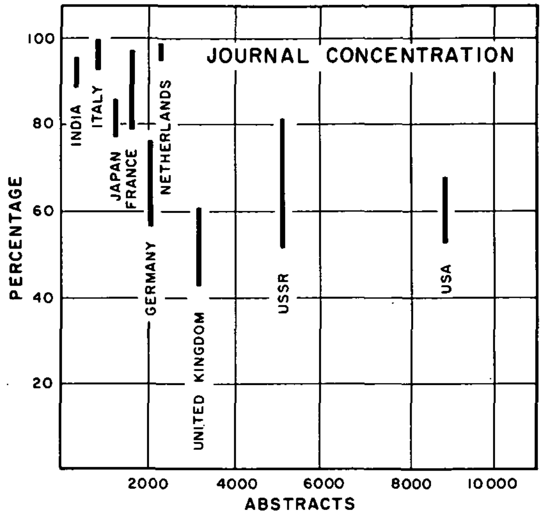


FIGURE XIV.27 Degree of concentration of journals published in various countries.[46] The abscissa of each vertical line segment is the number of abstracts in *Physics Abstracts* 1964 from journals published in the country shown. The ordinate of the lower (upper) end of the segment is the percentage of these abstracts contained in the five (ten) most productive journals published in this country. It should be remarked that the figures for Japan omit most Japanese-language journals.

is already such a multitude of journals that it is unreasonable to hope that a new journal devoted to some particular specialty will capture more than a fraction of the significant work in this specialty. For this reason, and because different specialties are becoming more and more interrelated, there is a growing danger that physicists will miss more and more of the papers that might be of importance to them and that reliance on specialist journals will worsen the problem.

A possible solution has been receiving serious consideration at the American Institute of Physics.[74] An editor or editorial board oriented toward any one of a large number (perhaps 50–100) of special subfields of physics might select, from articles scheduled for publication in any of several score key physics journals, those articles likely to be of interest to specialists in this subfield. These articles would then be gathered together and printed in a single issue, with suitable referencing of the archival journals from which they were taken. Such groupings—called "user journals"—could be circulated inexpensively to individuals interested in each specialty; the same article might well appear simultaneously in the user journals of two or more specialties, if it proved relevant to all of them. If such a system were put into practice, the original journals would retain their place as archival repositories for use in libraries and would continue to provide a useful publication outlet for authors; but the current-awareness role of journal scanning—which we have seen in Figures XIV.5 and XIV.7 to be very large—would be to a considerable extent taken over, and greatly improved, by the new user journals. Prepublication microform transmission of page proofs

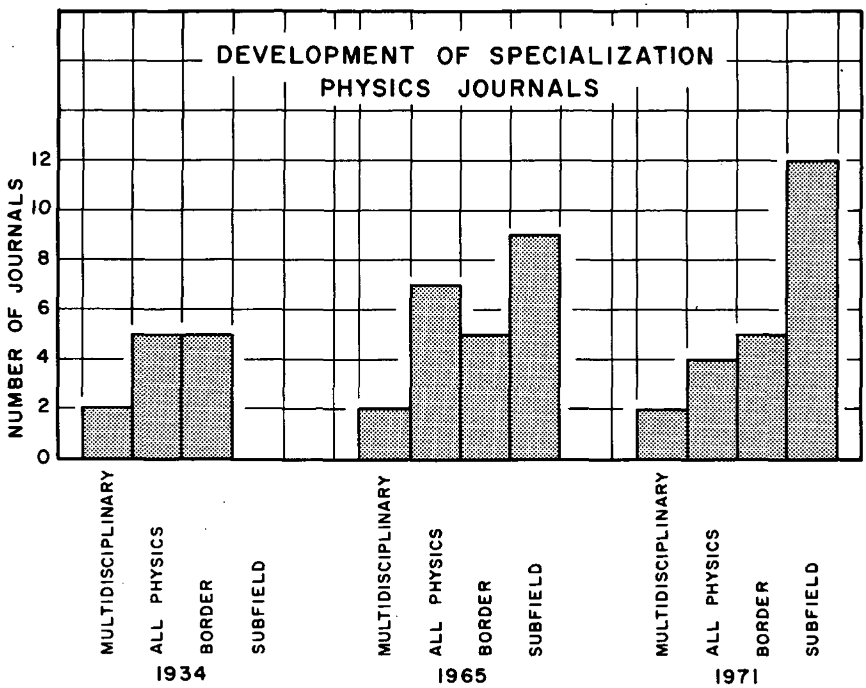


FIGURE XIV.28 Time development of the degree of specialization of the leading physics journals. Journals are classified as multidisciplinary (that is, covering all of physics and one or more other disciplines as well), all physics (covering all areas of physics), border (covering areas such as instrumentation or physical chemistry that form interfaces with physics), or, subfield (devoted to a subfield entirely within physics). The 1934 data refer to the 12 journals that accounted for half of the entries in *Physics Abstracts* in that year; the 1965 and 1971 data refer to the same set of 23 journals, which accounted for half of the entries in *Physics Abstracts* in 1965, but differ in that physics journals that have subdivided by 1971 have been transferred as units to the subfield column.

and photo-offset printing could make a very fast service possible. Journals with page-charge or other subsidy of prerun costs could cooperate in such a venture with no financial difficulty; for other journals, suitable royalty arrangements would be necessary. The whole scheme would integrate nicely with some of the secondary services discussed at the end of Section 3.6, with attendant overall economies.

Let us look finally at the intellectual relationship of different journals to each other as revealed in the patterns of citations. Because of both geographical provincialism and the specialization of journals, the typical scientist who publishes in a given journal will pay more attention to certain journals than to others, and these preferences will be reflected in the papers he cites. The degree of relatedness of journals through their author-readers is revealed in the "citation matrix," a preliminary example of which[61] is shown in Figure XIV.29; the entries give the fraction of the citations made by articles of a given journal (head of a column) that are to articles that ap-

REFERENCES (From →; To ↓)	Phys. Rev.	Proc. Phys. Soc.	Phys. Rev. Letters	J. Appl. Phys.	Sov. Phys. - JETP	Physica	Nuovo Cimento	Zeit. Physik	Progr. Theor. Phys.	Sov. Phys.-Sol. State	Can. J. Phys.	Czech. J. Phys.	Phys. Fluids	J. Phys. Soc. Japan	Proc. Roy. Soc.	J. Chem. Phys.	Can. J. Chem.	J. Chem. Soc.	J. Phys. Chem.	J. Am. Chem. Soc.
Phys. Rev.	47.2	34.1	28.4	14.5	18.5	15.8	25.0	19.7	29.8	12.8	15.1	12.3	8.7	15.4	6.9	12.8			1.3	
Proc. Phys. Soc.	2.0	9.4	1.2	2.4		4.3	1.0	2.5	1.1	1.2	2.0	2.0		3.7	2.9					
Phys. Rev. Letters	12.6	1.6	29.5	1.8	2.5	1.7	14.4	4.3	13.7		2.6	1.0	2.0							
J. Appl. Phys.	1.3	2.4	1.8	23.0						2.1	1.4	3.5	3.4	2.6	1.0					
Sov. Phys. - JETP	2.8		2.6	1.3	32.0		3.6	2.2	1.1	8.3		2.5			1.0					
Physica	1.1					21.5			1.7		2.6	1.5	1.2	2.2	1.3					
Nuovo Cimento	4.0	1.6	4.5		3.1		21.4		8.4			2.0		1.6						
Zeit. Physik		3.1				3.0		20.4		1.0		2.5			1.4					
Progr. Theor. Phys.	1.5						3.7		25.7											
Sov. Phys. -Sol. State										23.8		1.8								
Can. J. Phys.											9.1									
Czech. J. Phys.												12.5								
Phys. Fluids													19.5							
J. Phys. Soc. Japan												1.0		16.2						
Proc. Roy. Soc.	1.8	6.2	1.1	2.7		4.3	1.7	2.5	2.0	1.0	4.8	4.0	3.3		14.7	3.3	1.1	1.2	2.1	
J. Chem. Phys.		3.9	1.1	2.1		5.0					2.1	3.0	13.1	5.4	6.5	33.4	8.4	1.1	7.9	5.4
Can. J. Chem.																	12.3	1.3	1.1	
J. Chem. Soc.															2.7		10.2	25.4	3.7	8.0
J. Phys. Chem.															1.3	1.4	3.0		12.8	2.4
J. Amer. Chem. Soc.						2.0									3.8	6.3	17.6	19.4	22.2	39.2

FIGURE XIV.29 Partial reference matrix of a family of journals.[61] Each entry represents the percent of references in the journal represented by the column that are to articles in the journal represented by the row.

peared in each of various other journals (rows). The concentration onto the diagonal is conspicuous, although the diagonal element is by no means always the largest element of a column, since larger journals are naturally favored. One can remove the influence of size and get a quality-type measure of the relevance of the material published in one journal to that published in another by dividing the number of times the first journal is cited in the second by the total bulk of the first journal. Figure XIV.30 shows a typical result for a solid-state section of the *Physical Review.* Comparison with the "Section K" part of Appendix XIV.A reveals some interesting differences, which may reflect either differences in the quality of different journals or pigheaded provincialism of American authors.

4.4 PRIMARY PUBLICATION IN BOOKS

Counts of entries in the "Conference Index" of *Physics Abstracts* for 1969 reveal that a little over 7 percent of all papers abstracted were in proceed-

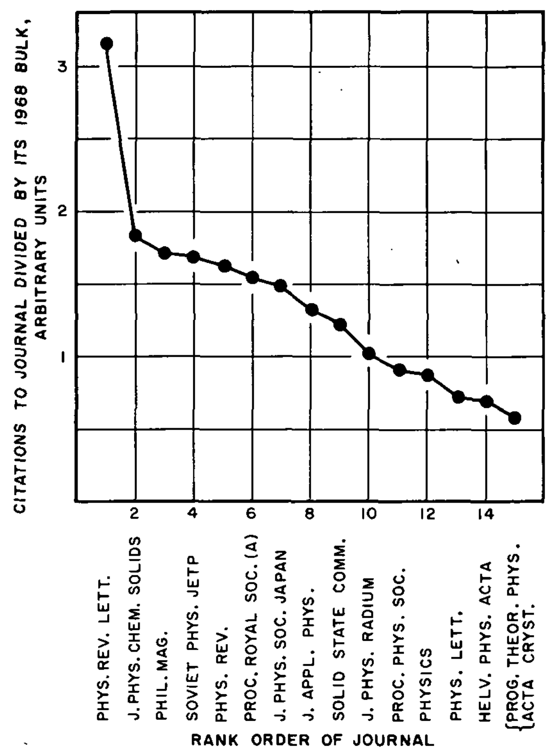


FIGURE XIV.30 Relative frequency of citation of different journals in a sample of Section 2 (the first of the solid-state sections) of the *Physical Review* in 1968. For each cited journal, the number of citations to this journal was divided by the current bulk of this journal to give a number roughly proportional to the probability for a given paper in the given journal to be cited. The journal with the highest value for this probability (*Physical Review Letters*) was assigned abscissa 1, that with the second highest probability (*Journal of the Physics and Chemistry of Solids*), abscissa 2, etc. Note that if the two solid-state sections of the *Physical Review,* which can be separately purchased by members though not by nonmembers, were treated as a separate journal, this journal would probably have abscissa 1 or 2 instead of 5.

ings of conferences. Almost three fourths of these appeared in regular journals, but the remainder (about 1.9 percent of the year's entries, or 2.3 percent of the research papers) were published in books. It is to these articles, plus a few items of original research published in collections of the *Festschrift* type, that we now turn our attention.

For many years, these papers were very inadequately covered by *Physics Abstracts,* but the coverage seems now to be fairly good, so it is probably safe to accept the figures just quoted as giving a fair estimate of the percentage of research papers that appear in books. (Some conferences publish their proceedings in a regular journal but give buyers the option of purchasing them in a separate binding; this we do not consider as book publication.) Although there are far fewer data readily at hand on this type of publication, we shall try to make rough estimates for it of some of the same characteristics we have discussed in the preceding subsections for journals.

4.4.1 *Bulk, Price, and Circulation*

As we have just noted, the number of research articles in physics published in books seems currently to be of the order of 2 percent to 2.5 percent of the total of such articles, that is, of the order of a thousand a year. This output appears in approximately 50 books, priced usually in the range of 5–12¢ per equivalent kiloword, that is, a range similar to that which obtains for journals of commercial publishers (see Figure XIV.19). Circulations can only be guessed, but one thing at least is likely: Besides library circulations that are probably only slightly less than those of journals with a comparable price per kiloword, most conference proceedings are purchased—often through registration fees—by many or all of the individual participants in the conference. These people, to whom the material is thus available, comprise a portion, but usually only a minor portion, of the population to whom it is of high interest.

4.4.2 *Time Lags*

The interval between the deadline for submission of manuscripts and the appearance of the proceedings of a conference or a *Festschrift* varies widely but is usually rather longer than the 5–7-month lag typical of research journals in physics (Figure XIV.25); delays of over two years are by no means unknown. Conferences that occur regularly (for example, every year) can often establish a regular production schedule and by continuing experience minimize the various sources of delay. But one-shot conferences, often dependent on inexperienced editors or program committees, are apt to run into all sorts of unanticipated delays.

4.4.3 *Language Distribution*

The predominance of English seems to be even more marked for conference proceedings than for journal articles. Conferences published in books are often international, and since the oral communications are usually in English, the language most likely to be understood by a majority of conference participants, the written papers are usually in English also.

4.4.4 *Quality and Refereeing*

Most conferences that publish their proceedings have a program committee that referees papers submitted; papers in *Festschrift* volumes usually appear by invitation only. Thus one would expect quality standards to be similar to those in journals, which we have discussed in Section 4.3. However, there are undoubtedly differences in the refereeing criteria. Perhaps more serious is the effect of time and space limitations on quality: Papers in conference proceedings are usually strictly limited in length, they must be submitted by a definite deadline, and, as noted, they are apt to be delayed in publication. Thus authors are likely to publish their best material elsewhere. The subjective judgment of many physicists is, in fact, that the average quality of papers in conference proceedings is lower than that of papers in the better journals.

Nevertheless, citation statistics are not altogether unfavorable to papers published in books; for example, in a sample of all four sections of the 1971 *Physical Review* the ratio of citations to papers in books devoted to conference proceedings, or in *Festschriften,* to the citations of journal articles, was roughly 3.25:4 percent; this is rather larger than the ratio of the number of such papers in *Physics Abstracts* to the number of journal articles, which we estimated to be less than 2.5 percent. However, it must be borne in mind that the accessibility of the conference books to the authors would be higher than that of many obscure or foreign language journals that contribute to the entries in *Physics Abstracts.* There is fragmentary evidence that conference papers are rather less likely to be cited than those in leading journals; for example, in a study made some years ago (Herring, 1966, unpublished data) it was found that papers in the annual Conference on Magnetism and Magnetic Materials (published in the *Journal of Applied Physics,* not in a book) had only a fraction as many entries in the *Science Citation Index* as papers in the same field in *Physical Review.* More thorough studies are needed; in particular, it would be interesting to check the reality of some effects that seemed to appear in the present counts from the 1971 *Physical Review,* namely that atomic and molecular papers and theoretical particle-physics papers cite conference books very little.

An obvious recommendation with which to close is that *wherever possi-*

ble, the managers of conferences should arrange to publish the proceedings (if publication is deemed necessary) in a journal of wide circulation, rather than in a special book. The proceedings issue should, of course, be available to individual participants at a low price. (We shall return to this matter briefly in Section 6.1.)

4.5 CHARACTERISTICS OF RESEARCH PAPERS IN PHYSICS

We turn now to research papers—the intellectual and sociological aspects of the work they describe and their quality and style. Much of the material to be discussed in the following paragraphs has been obtained from a sampling of papers listed in *Physics Abstracts,* 1969, so it is appropriate to begin with a description of our sampling experiment. Two physicists (Robert W. Keyes and Conyers Herring) acted as samplers, one looking at all entries in 13 of the year's 26 issues whose numbers ended in 5, the other looking at all entries in 16 of the issues whose numbers ended in 0. Entries not in the category of "published research papers" (as defined in the notes to Figure XIV.11 and Tables XIV.10a and XIV.10b were discarded. Each paper not discarded was classified into one of the ten subfields of physics corresponding to the panels of the Physics Survey, or into a "miscellaneous physics" category, or (in rare cases) as lying outside physics. These judgments were made subjectively by each sampler but following the guidelines described in the notes to Table XIV.10. Each paper was also classified as theoretical, if no new experimental data or techniques were presented, or as experimental. Each paper was further assigned to a country of origin according to the location of the first-named institution in the byline, except when the institution was operated by an international organization, in which case the assignment "international" was used. If no institution was named, the location of the publisher of the paper was recorded and used as described in the notes to Table XIV.10. For papers from U.S. institutions, the institutions were further classified as academic, industrial, government in-house, federally funded (that is, national laboratories), or other.

Figure XIV.31 gives a quick overview of the distribution of published research papers over the subfields and over the nations or regions. More detailed numbers appear in Tables XIV.10a and XIV.10b, which are separated to show the extent of the agreement in the data collected by the two samplers. We call attention to the following highlights:

1. Theoretical work in the United States is more prolific than is the case for Western Europe and the Soviet Union in all subfields of physics; in regard to experimental work, the United States is again more prolific in most

TABLE XIV.10a Distribution of a Sample of 2032 Published Research Papers taken from *Physics Abstracts* 1969 over Subfields of Physics and Geographic Region of Performing Institution[a]

Location of Performing Institution	Total All Subfields		A&R		AME		EP		NP		P&F		CM		E&P		PB		Op		Ac		Misc.	
	Exp.	Theor.	Exp.	Theor.	Exp.	Theor.	Exp.	Theor.	Exp.	Theor.	Exp.	Theor.	Exp.	Theor.	Exp.	Theor.	Exp.	Theor.	Exp.	Theor.	Exp.	Theor.	Exp.	Theor.
U.S.A.	374	320	2	10	30	43	16	38	27	26	12	37	196	88	48	36	3	0	26	16	2	6	12	20
Other No. Amer.	34	42	1	1	5	5	0	1	5	3	2	6	10	15	4	6	0	0	4	4	1	0	2	1
So. Amer.	3	4	0	0	1	1	0	1	0	1	0	0	2	1	0	0	0	0	0	0	0	0	0	0
U.K.	103	71	0	2	9	13	2	10	2	2	10	8	60	20	12	7	1	0	5	2	1	3	1	4
France	74	39	0	3	11	6	2	2	7	2	3	5	31	15	8	1	0	0	8	0	2	1	2	4
W. Germany	78	49	0	0	6	6	5	6	9	11	2	3	43	13	3	3	0	0	8	5	0	1	2	1
Other W. Europe	83	55	0	2	5	2	7	14	9	4	1	7	46	16	7	3	0	0	4	2	0	1	4	4
U.S.S.R.	235	173	1	0	12	18	8	14	18	15	21	25	120	54	23	24	0	0	25	15	1	0	6	8
Other E. Europe	30	45	0	1	1	4	1	7	4	4	0	3	18	20	3	3	0	0	3	1	0	0	0	2
Africa	4	2	0	0	0	0	0	0	3	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0
Japan	58	38	0	0	5	4	0	6	4	3	4	6	34	14	4	3	0	0	5	1	1	0	1	1
India	35	28	0	1	4	5	0	5	2	1	1	5	21	9	3	1	0	0	3	0	0	0	1	1
Other Asia	11	14	0	0	0	3	0	4	4	1	2	2	5	1	0	0	0	0	0	1	0	0	0	2
Australia & N.Z.	18	12	0	1	2	1	0	0	2	1	1	3	8	4	4	2	0	0	1	0	0	0	0	0
Total	1140	892	4	21	91	111	41	108	96	75	59[b]	110[b]	595	271	119	89	4	0	92	47	8	12	31	48

[a] Total entries scanned, 2492, of which 2032 were published research papers.
[b] Plasma = 33 experimental and 52 theoretical; fluids = 26 experimental and 58 theoretical.

TABLE XIV.10b Distribution of a Sample of 2576 Published Research Papers taken from *Physics Abstracts* 1969 (#855–#870) over Subfields of Physics and Geographic Region of Performing Institution[a]

Location of Performing Institution	Total All Subfields		A&R		AME		EP		NP		P&F		CM		E&P		PB		Op		Ac		Misc.	
	Exp.	Theor.	Exp.	Theor.	Exp.	Theor.	Exp.	Theor.	Exp.	Theor.	Exp.	Theor.	Exp.	Theor.	Exp.	Theor.	Exp.	Theor.	Exp.	Theor.	Exp.	Theor.	Exp.	Theor.
U.S.A.	436	359	2	11	44	68	15	57	42	20	19	45	234	83	29	34	23	0	15	10	7	8	6	23
Other No. Amer.	52	28	0	1	7	4	1	4	14	3	2	3	21	8	2	0	2	0	1	1	0	0	2	4
So. Amer.	1	2	0	0	0	1	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0
U.K.	104	87	0	2	17	13	0	14	8	5	7	12	50	26	6	4	4	0	5	2	0	0	7	9
France	74	41	0	2	10	6	0	4	2	0	1	4	50	14	7	3	1	0	2	2	1	1	0	5
W. Germany	95	45	0	0	19	11	2	2	20	7	5	5	40	13	0	2	1	0	3	0	4	0	1	5
Other W. Europe	110	71	0	4	12	6	1	15	24	11	4	6	50	16	6	3	1	0	5	2	0	1	1	7
U.S.S.R.	264	173	2	6	25	23	9	13	14	16	14	26	151	36	22	19	2	0	13	10	5	6	7	18
Other E. Europe	50	51	0	2	5	4	0	7	8	5	0	9	29	12	3	1	2	0	1	0	1	0	1	11
Africa	6	5	0	0	2	0	0	1	2	1	0	1	2	0	0	1	0	0	0	0	0	0	0	1
Japan	100	50	0	0	5	4	0	11	14	3	7	9	69	13	4	2	0	0	1	0	0	3	0	5
India	46	47	2	2	10	6	0	5	4	3	1	11	24	12	2	0	0	0	0	2	0	2	3	4
Other Asia	8	10	0	0	1	1	0	2	3	1	0	1	4	3	0	1	0	0	0	0	0	0	0	1
Australia & N.Z.	23	8	2	0	0	0	0	0	3	3	0	0	9	4	0	0	2	0	2	0	2	0	3	1
Unknown	137	93	7	10	11	6	1	12	10	10	5	11	56	15	21	10	5	0	13	7	3	1	5	11
Total	1070	1506	15	40	167	152	29	148	168	88	65	144	797	255	102	80	43	0	61	36	23	22	36	105

[a] Total entries scanned, 2980, of which 2576 were published research papers. First figure in each section refers to experimental papers; the second, to purely theoretical papers.

A&R = Astrophysics and relativity
AME = Atomic, molecular, and electron physics
EP = Elementary-particle physics
NP = Nuclear physics
P&F = Plasma physics and the physics of fluids
CM = Condensed matter
E&P = Earth and planetary physics
PB = Physics in biology
Op = Optics
Ac = Acoustics

The numbers in Table XIV.10 (parts a and b) refer to published research papers, that is, presentations of new research results in a periodical or a book available on the open market. Among the items excluded from this category are theses and reports available only from a research institution, agency, or clearinghouse; patents; reviews; popularizations; abstracts unaccompanied by full text. Letters in regular letter journals are included.

Papers that appeared to be clearly closer in content to chemistry, astronomy (with the exception of astrophysics and relativity), mathematics, and like disciplines than to physics were excluded. These papers constituted approximately 6 percent of the total number of papers scanned.

Assignment of a paper to a country or region was made on the basis of the institutional affiliation of the first author, if this information was given. When such information was not given in the abstract, the assignment was made on the basis of the authors' names and the place of publication for cases in which such assignment could be regarded as about 95 percent certain (for example, authors with Russian names publishing in a Russian-language journal); other cases were listed as "Unknown" in Table XIV.10b. However, in Table XIV.10a and Figure XIV.31, these unknown cases were distributed among the various regions based on the known publication patterns of the principal journals involved, which often have a mainly national clientele and sometimes a very cosmopolitan one. The estimated distribution for Table XIV.10a was:

	A&R		AME		EP		NP		P&F		CM		E&P		PB		Op		Ac		Misc.	
	Exp.	Theor.	Exp.	Theor.	Exp.	Theor.	Exp.	Theor.	Exp.	Theor.	Exp.	Theor.	Exp.	Theor.	Exp.	Theor.	Exp.	Theor.	Exp.	Theor.	Exp.	Theor.
U.S.A.	2	3	7	2	1	0	2	3	3	2	24	7	10	4	4		9	2	1	0	1	3
W. Eur.	3	4	3	3	0	0	5	2	0	4	25	5	7	4	1		4	5	2	1	4	7
U.S.S.R.	0	1	0	0	0	0	1	0	1	0	2	1	1	1	0		0	0	0	0	0	0
Japan	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0		0	0	0	0	0	0
Other	2	2	1	1	1	0	0	1	1	2	4	2	2	1	0		0	0	0	0	0	1

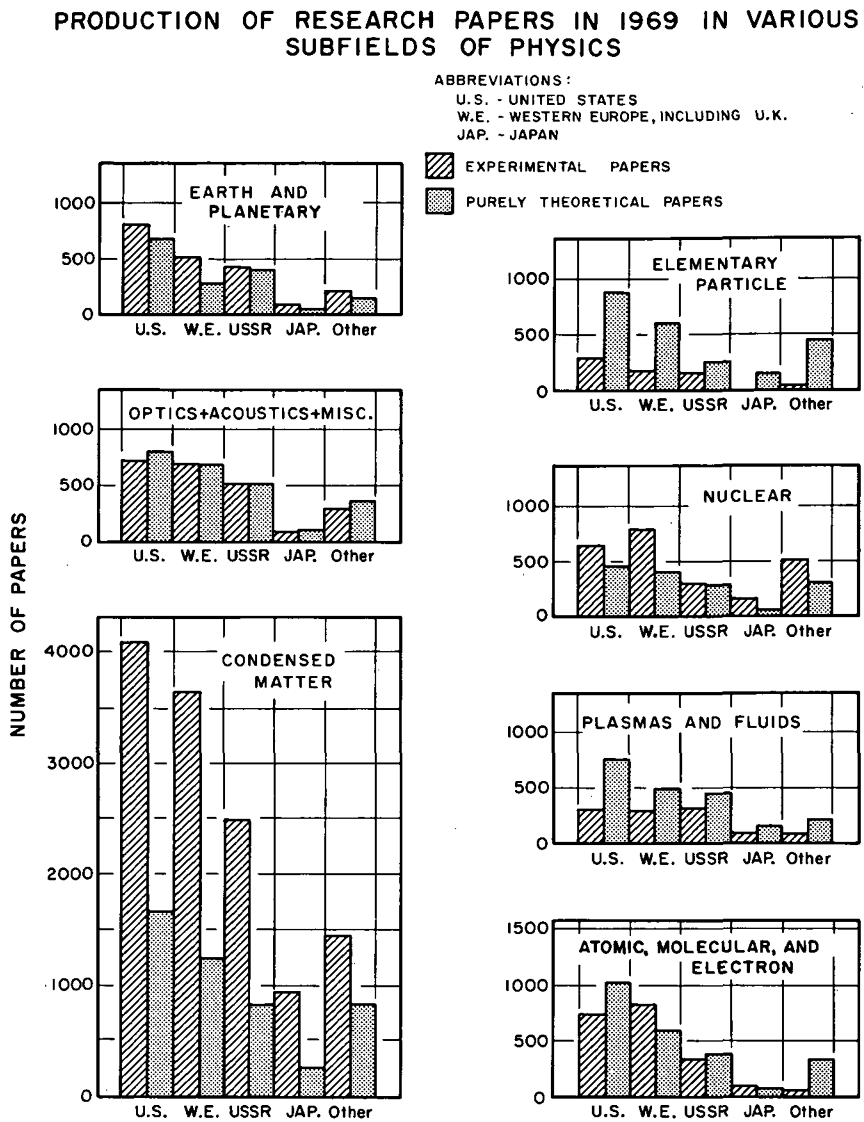


FIGURE XIV.31 Production of research papers in 1969 in various subfields of physics by institutions in various countries or regions. (Data are based on Table XIV.10, scaled up to total annual production.)

subfields but is slightly behind Western Europe in nuclear physics and in atomic, molecular, and electron physics; it also may be slightly behind the Soviet Union in plasmas and fluids.

2. Theoretical papers greatly outnumber experimental ones in elementary-particle physics and relativistic astrophysics, and somewhat outnumber them in plasmas and fluids and, in some cases, in atomic, molecular, and electron physics. In all other subfields, experimental papers are preponderant, and in biophysics there are essentially no purely theoretical papers.

3. As one would expect from the costliness of the equipment required, experimental work in elementary-particle physics is almost entirely confined to the United States, the Soviet Union, and Western Europe. There is a similar, though less pronounced, tendency toward both experimental and theoretical work in earth and planetary physics in these countries.

4. Though the data are rather poor because of the sketchy coverage of *Physics Abstracts* in this subfield, work in biophysics appears to be overwhelmingly concentrated in the United States.

The distribution of published papers from U.S. institutions over subfields and kinds of institutions appears in Table XIV.11 and Figure XIV.32. We note:

5. About 57 percent of U.S. papers originate in academic institutions, 21 percent in industrial laboratories, 11 percent in federally funded research

TABLE XIV.11 Distribution of Sample of Papers from U.S. Institutions in 23 Issues of *Physics Abstracts* 1969 by Type of Performing Institution[a]

Performing Institution	Total		Subfield Breakdown										
	Exp.	Theor.	A&R	AME	EP	NP	P&F	CM	E&P	PB	Op	Ac	Misc.
Academic	308	372	17	103	79	53	59	236	61	12	26	8	26
Industrial	153	90	0	17	2	5	16	152	18	1	20	4	9
Government laboratory	67	40	2	10	3	6	6	39	27	0	5	4	5
Federally funded research and development center[b]	95	33	0	8	15	22	8	52	9	1	8	0	6
Other	14	11	1	4	2	1	0	10	2	1	2	1	1

[a] Total entries scanned were 4160. (Note that this total is less than in Tables XIV.10a and XIV.10b.) Experimental and theoretical papers are not separated except in the totals for all subfields taken together.

[b] As defined in Reference 75, pp. 97–99.

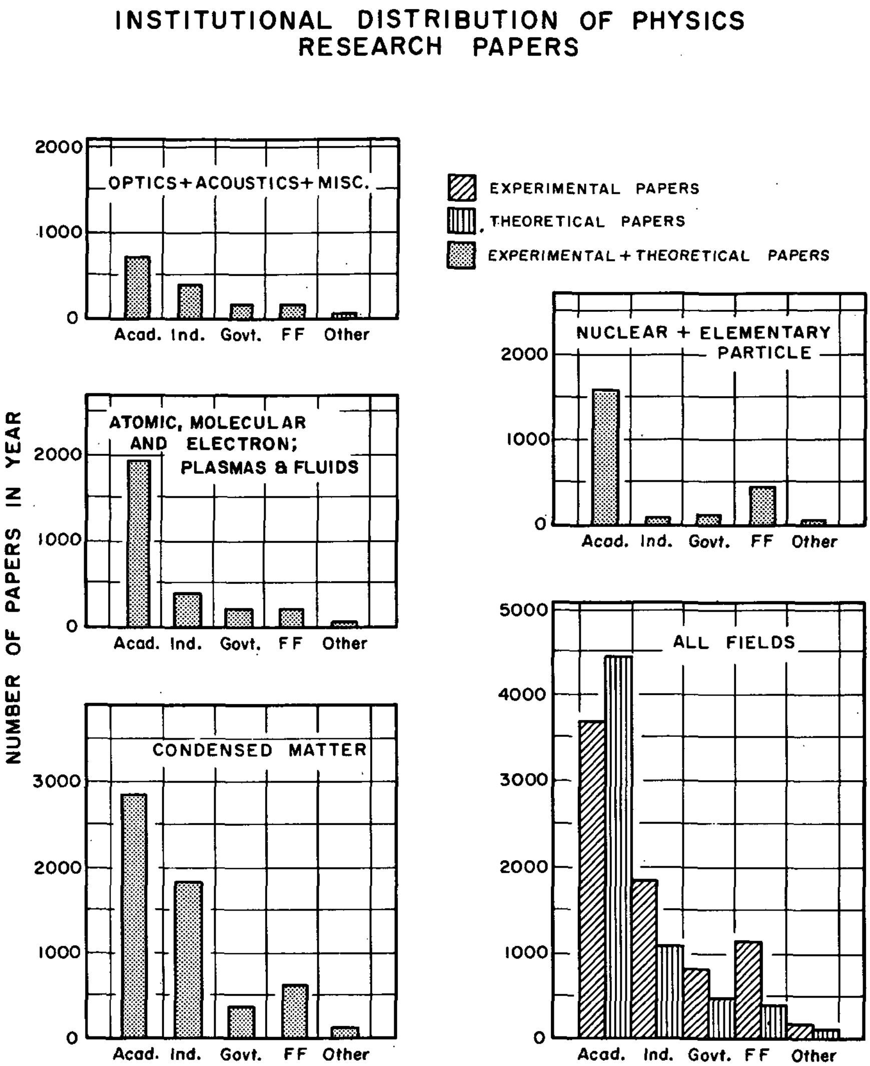


FIGURE XIV.32 Distribution of production of research papers in various subfields of physics over different types of U.S. institutions. (Theoretical and experimental papers are grouped together except for the total for all subfields.) Data are based on Table XIV.11.

centers, 9 percent in government in-house laboratories, and 2 percent elsewhere.

6. Slightly more of the work in academic institutions is theoretical than experimental; all the other types of institutions produce more experimental

than theoretical papers. Although not shown in Table XIV.11, the difference is especially marked in elementary-particle physics, in which work in universities is overwhelmingly theoretical and that in the federally funded research centers predominantly experimental; most of the experimental work is, in fact, done in the latter institutions.

7. About 63 percent of the papers from industrial laboratories are in condensed matter, yet even in this subfield, academic institutions produce half again as many papers as industrial. A similar distribution between these two types of institutions obtains in optics, acoustics, and miscellaneous physics.

Let us now turn from these gross statistics to some more subtle characteristics of research papers. One is length. The length of a paper is influenced by many factors: Authors and editors are both interested in striking a good compromise between the extremes of incomprehensible terseness and repellent prolixity, but authors are apt to be lazy about reducing their verbiage, while editors are apt to worry about their budgets for composition; conference papers are usually subject to ceilings on length; authors in highly competitive fields are apt to publish each new little bit of progress as rapidly as they can, rather than waiting to collect material for a complete story. Whatever the reasons, however, statistics (Figure XIV.33) show that papers in the *Physical Review* are longer, on the average, than they were one or a few decades ago. This finding suggests that papers in general have been getting longer, but the point cannot be considered established until studies of the entire literature are made, since the distribution of long and short articles between the *Physical Review* and other journals could conceivably change with time. Such studies would have to take ac-

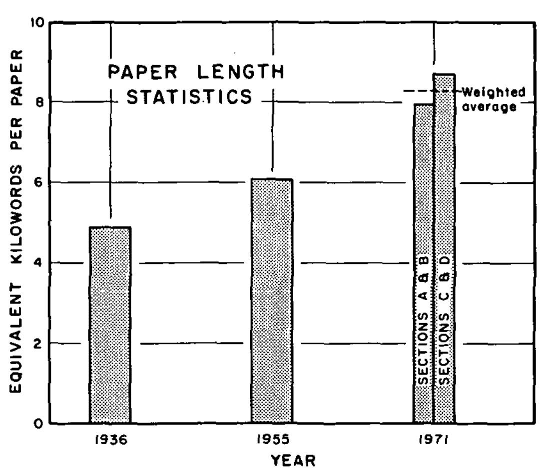


FIGURE XIV.33 Increase, over the decades, in the average length of full papers (excluding letters, comments, and the like) in the *Physical Review*. (Equivalent kilowords means the number of words that would occupy the space used by a paper if this space were set in solid text.)

count of variations in typography and page size from journal to journal and their change with time. (The tabulations in ICSU-AB[45,46] give extensive data on lengths of papers in pages, but without consideration of those latter factors. Anthony *et al.*[38] have given figures on average equivalent words for 1964 journal articles published in different countries.)

Another characteristic of papers, which is even more directly related to the changing conditions of research, is the multiplicity of authorship. The upper part of Figure XIV.34 shows, both for full papers and for letters, how the average number of authors of a paper (letter) has changed over the decades in the journals of the American Physical Society. The trend toward increasing numbers of authors is clear. However, the trend has not been felt equally in all subfields of physics; as the lower part of Figure XIV.34 shows, papers in nuclear physics have, on the average, over one and one half times as many authors as those in atomic, molecular, and fluid physics; intermediate numbers occur for solid-state and elementary-particle papers. There are probably two main effects: Experimental work using large machines tends to produce multiauthored papers; theoretical papers tend to have few authors. In elementary-particle physics, the former effect is strong, but the great numerical preponderance of theoretical papers (Tables XIV.10a and XIV.10b and Figure XIV.31) smothers it. Single-author papers have become a rarity in nuclear physics (8 percent) but are still fairly common in the other three *Physical Review* sections (an average of 35 percent). Letters, for some reason, tend to have more authors than papers. Vlachy[63,76] has pointed out further differences, for example, among different countries and among different disciplines.

Still another measurable characteristic of papers is the number and distribution of the references they cite. Data on these are laborious to accumulate by hand, as one cannot simply count the number of footnotes. Often one footnote will refer to several papers, and often a footnote is merely a remark, not a reference; customs in such matters as these may change with time and may vary from journal to journal. However, computerized records such as the TIP file[61] or the AIP SPIN tapes[42] could be very useful for citation studies. We have made a few counts for samples of full-length papers in 1936 and 1971 issues of the *Physical Review,* which show:

1. The average number of journal articles cited per paper has increased from about 12.5 in 1936 to about 17.5 in 1971. But as we have seen in Figure XIV.33, the 1971 papers are on the average nearly 1.7 times as long as those of 1936. So the number of references has increased more slowly than the length.

2. Theoretical and experimental papers do not differ greatly in numbers

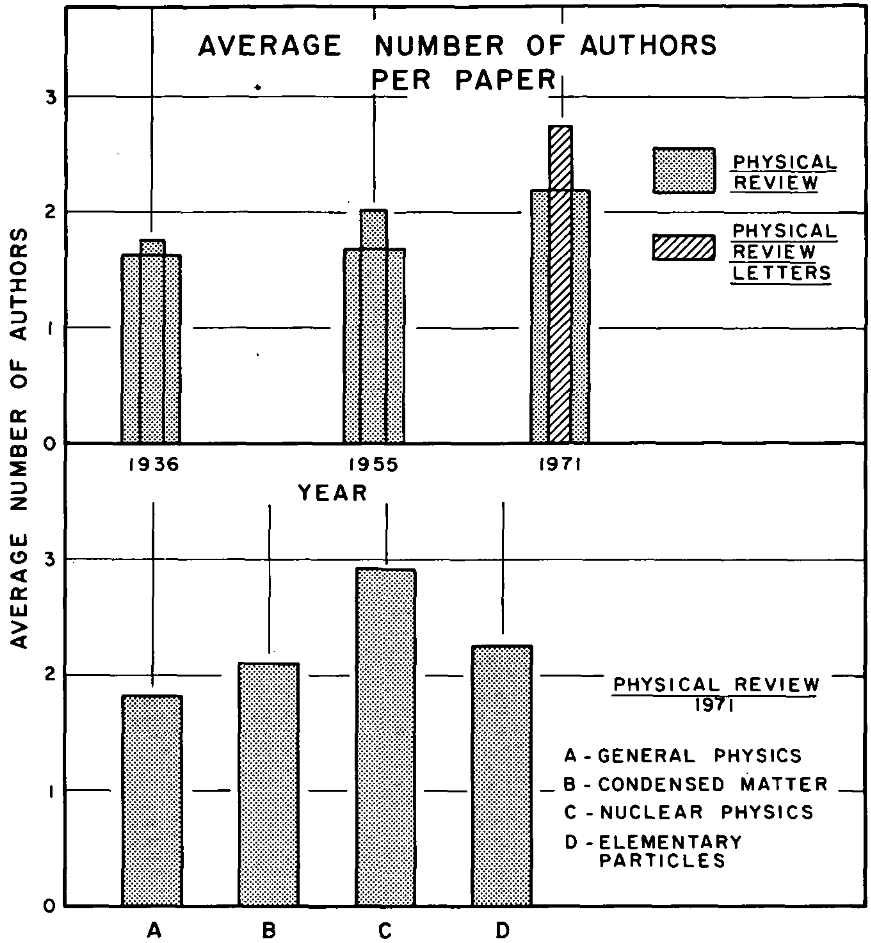


FIGURE XIV.34 Average number of authors per paper in the *Physical Review* or *Physical Review Letters. Top:* Growth in numbers of authors over the decades (1971 figures are for the total of all sections of the *Physical Review*). *Bottom:* Variation in average numbers of authors among physics subfields in 1971 as exemplified by the comparison of *Physical Review* Sections A (general physics, principally atomic, molecular, and fluids), B (solid-state physics), C (nuclear physics), and D (particles and fields).

of references to journal articles, except in particle physics, where the number is low (about 10.5) for the theoretical papers and over twice as high for the experimental.

3. Citations to unpublished theses and reports and to research articles in books were much less common in 1936 than today. (See Section 4.8.)

The distribution of citations in age depends on the rate of obsolescence of the cited papers and, doubtless, on the psychology of the citing authors. Burton and Kebler[77] have studied the distribution of citations in the literature of a number of fields, including physics. Price[78] has made an interesting analysis of the distribution of citations in all areas of science together. Most useful for us, however, are some unpublished data of Schawlow (1967) on citations in papers of the *Physical Review*. These are summarized in Figure XIV.35, from which we note:

4. In most subfields of physics, the frequency of citation of past literature decreases exponentially with increasing age at a rate corresponding to a factor two in about 3.5 years. But in elementary-particle physics, the decrease is significantly faster, especially for the most recent years (perhaps an "immediacy effect"–see Price[79]).

5. This decrease is compounded in comparable amounts by the exponential growth of the literature capable of being cited (Figure XIV.10) and the obsolescence of this literature. This obsolescence is roughly measured by the upper curve in Figure XIV.35, though properly this curve should be corrected for time variation in the quality of coverage of the literature by *Physics Abstracts*. The upper part of this curve, which Figure XIV.10 suggests should be the more reliable, gives an obsolescence half-life of about eight years. We shall include these data in a more complete analysis of obsolescence in Section 4.8.

(These data are, of course, also relevant to patterns of use of the literature, a topic to which we shall return in Section 4.8.)

Statistics such as these raise another question about research publications: To what extent is the distribution of work currently being published governed by fads of the moment? There are two competing effects. People like to rush in on what is new and exciting; on the other hand, many workers, especially beginners in research, are only too glad to find a subject of interest that is currently being ignored by others. There is a smattering of evidence that the second effect is quite strong. For example, one might expect the dominance of the first effect to be greatest in a fast-publication, high-prestige journal such as *Physical Review Letters;* yet we have found, in a sampling of a number of the 2000 or so fine specialties into which solid-state physics can be divided, that the number of these on which at least one letter has appeared in the last five years is only modestly below what would be expected for a random distribution weighted by the amount of work in earlier decades. In other words, no major proportion of the topics on which research was done in the 1940's and 1950's has been neglected by the workers who publish in *Physical Review Letters;* this statement should be true *a fortiori* for the literature as a whole.

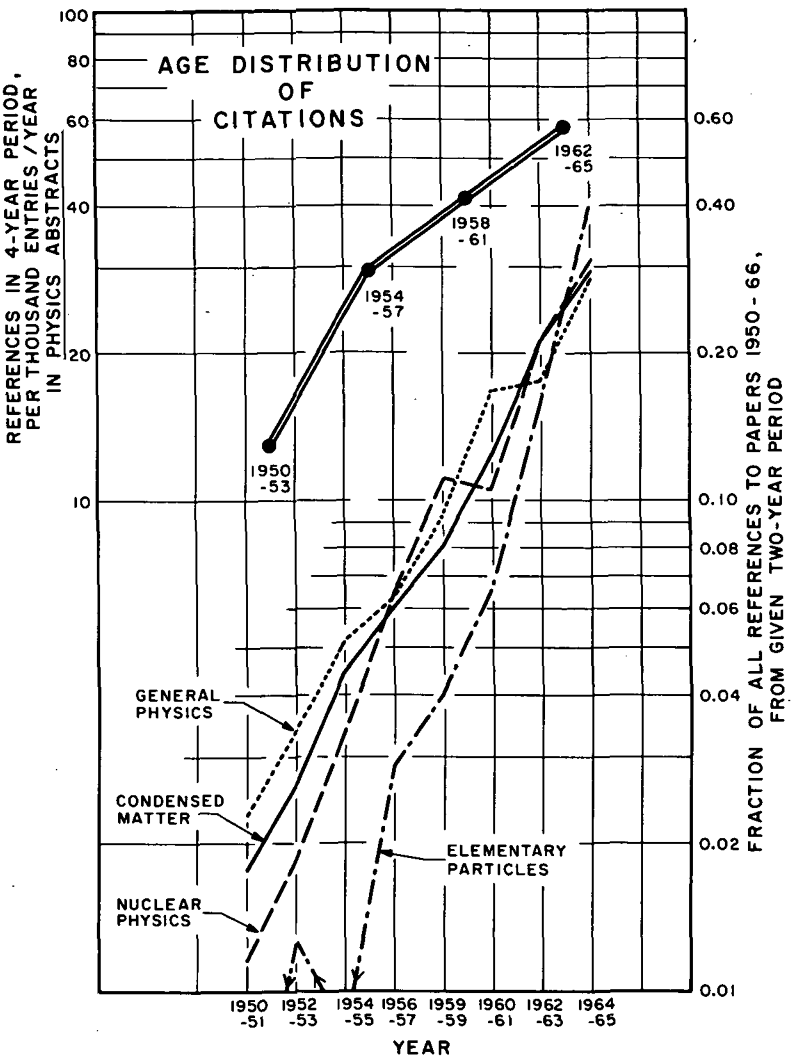


FIGURE XIV.35 Age distribution of citations in a sample of papers from the *Physical Review*, 1967. Lower curves (scale at right): fraction of all references to papers dated 1950–1966 that were to papers published in each of eight successive two-year periods, for each of four parts of the *Physical Review*. (Issue 1, atoms, molecules, fluids, and general physics; issues 2 and 3, solid state; issue 4, nuclear physics; issue 5, particles and fields.) Data for papers published in 1966 are not shown, as most of the citing papers were submitted during 1966. Upper curve (scale at left): total references in all issues of the *Physical Review* divided by number of papers in *Physics Abstracts* for the corresponding year and summed over four-year intervals. [Source: A. L. Schawlow, 1967, unpublished data.]

4.6 PREPRINTS, REPORTS, THESES, AND OTHER CONSIDERATIONS

We have seen in Section 3.2 (especially Figure XIV.11) that various kinds of documents that are unpublished—though often available to anyone on special order—constitute a sizable minority of the items deemed worthy of listing in some of the leading abstract journals of physics. Specifically, reports and theses account for about 3 percent of *Physics Abstracts* and about 25 percent of *Nuclear Science Abstracts* (physics part). In addition, many (probably most) papers that are submitted for publication in journals are sent in preprint form by the authors to tens or even hundreds of their colleagues in other institutions, or are distributed within the authors' institutions as internal memoranda. Physicists tend more often to distribute preprints than do mathematicians, biologists, or chemists.[31]

Since the production of preprints, reports, and the like is done by many small organizations, and since their distribution is largely by individualized mailing lists or individual requests, it is much more difficult to get an adequate overall picture of the extent of circulation of such material than is the case for published journals. Overall statistics that may be available from some large distribution centers (National Technical Information Service, Defense Documentation Center, and the like) usually do not separate physics from other scientific and technological fields; inasmuch as engineers place much more reliance on reports than do scientists, such statistics may not give a correct picture for physics. Still, a few tidbits from these overall statistics may provide a helpful perspective; some of these, and many references to earlier literature, can be found in a discussion of the role of reports in communication[80]:

1. Announcements by the Federal Clearinghouse (now the National Technical Information Service) amounted, a few years ago, to about 30,000 reports a year; however, a much larger number of reports produced by government agencies and contractors are classified or of limited distribution.
2. Of the reports deposited with the Defense Documentation Center, 10 percent remained unrequested, 80 percent received from one to ten requests, and 10 percent received more than ten requests.
3. Nearly two thirds of all unclassified technical reports contain information worthy of publication, yet, for about one fifth of these reports with publishable material, none of this information actually gets published, at least for several years.[81]

For physics proper, the most detailed source of information about preprints and reports is a study by Libbey and Zaltman[18] of the role and distribution of such material in theoretical high-energy physics. These authors queried all known high-energy theorists in the world about the origination, reception, and use of unpublished documents and also queried selected larger institutions about their policies and experiences. Responses were received from 59 percent of the estimated 1722 high-energy theorists queried.

Let us look first at the gross bulk of circulation of unpublished documents. A question asking for the distribution of preprints of material destined for publication (probably about half of all the documents involved) was answered by 410 respondents (24 percent). For these:

4. The average number of copies sent out was 148; the average number of typewritten pages was 22. But fluctuations in these numbers were large, especially in the sizes of the distributions.

In the survey of larger institutions, mailings by the institution ranged from 50 to 800 (average 233) for 16 U.S. institutions, and from 25 to 300 (average 137) for 11 foreign institutions. Also:

5. Slightly over half of the preprints referred to in 4 above were sent out simultaneously with their submission for publication; about two thirds of the remainder were sent out before submission for publication, and one third after.

6. The average respondent received about 186 documents of all types per year. If we guess the average length of these to be the same as that of the preprints in 4, a rough estimate of $186 \times 22 \times 1722 = 7$ million typewritten pages of this material results. At an estimated runoff cost of 1.6¢ per page, distribution costs amount to about \$110,000 per year. (This estimate of runoff cost is based on the estimates of Libbey and Zaltman in their Section XI for costs of a centralized distribution service, which probably would be more efficient than the multitude of smaller-scale productions that are now involved.) For comparison, the total world investment (prerun and runoff) in publication of the estimated 2320 (1969) papers in high-energy theory can be estimated, from Tables XIV.10a and XIV.10b and Figure XIV.33, and the cost figures of Section 4.2, at approximately \$1.8 million, plus a profit figure for the part put out by commercial publishers. Thus *the investment in the present modest circulation of preprints and other unpublished documents in this field is small compared with that in publication* (*though it may be one third or one half the runoff part of the latter*).

Both in physics and in other fields of science there have been lively controversies in recent years over the desirability of centrally subsidized schemes for large-scale distribution of preprints.[82–84] The advantages adduced are its speed of communication and its selectivity, since, in theory, each document can be routed only to those recipients who have expressed an interest in the topic it addresses. The opponents of such schemes argue that in terms of runoff cost such schemes are relatively costly, and that the widespread distribution of unrefereed documents would result in scientists' being flooded with a great bulk of material of low quality, in which it would be difficult to identify the material worthy of perusal.

Fortunately, data of the type we have been discussing above are a helpful guide toward finding a commonsense middle ground. In high-energy theory, for example, we have seen that the present investment in *distribution* of unpublished documents—assuming that at least one copy of each has to be produced anyway—is small, so that this distribution can surely be justified if it adds even a modest percentage to the utility of the published channel. According to Figures XIV.4, XIV.5, and XIV.7, preprints supply 10–20 percent of useful ideas to research physicists; if these data are at all applicable to high-energy theory, and if the nuisances of receiving stacks of preprints are not serious enough to compensate for this, the expense is doubtless justified, since in such a fast-moving field the information contained in the preprints is of appreciably more value than it would be if one had to wait five or six months to see it in print (see Figure XIV.35). But expansion of the circulation of preprints by a large factor would probably not increase the number of *readers* of each preprint by a like factor. We have seen in 6 above that the average high-energy theorist already receives 186 preprints a year, surely many more than he can really read. What is needed is not an increase in absolute numbers—even a decrease might be desirable—but better aim in distribution. Newcomers to a field should have the same opportunity to receive preprints as "big shots," and distribution to those not really interested in the subject should be minimized.

These objectives, and also the objective of "bibliographic control" (SATCOM,[2] pp. 71–72), could be realized by a service that would announce titles of available preprints or other unpublished documents and mail them only to individuals requesting them by return postcard, or to those libraries that wish to maintain complete files of all preprints in some field and are willing to pay the full costs involved. Such a service would not need to interfere with the freedom of authors to send out preprints to their friends, but it would relieve them of the need to send out hundreds of (usually unwanted) copies. The service might well send, together with its announcements of new preprints available, announcements of former items now available in print, which libraries or individuals could then discard. Journals of course should disallow references to preprints appearing on the latter list.

4.7 QUALITY OF PUBLISHED RESEARCH IN DIFFERENT FIELDS

We come now to a question of great conceptual interest and practical importance: Can one, in any objective way, compare the quality of work in different fields or at different times? Some grounds for hope have emerged from an experiment we have conducted with the assistance of the Panel on the Physics of Condensed Matter, though the results are only preliminary. In this experiment, a random sample of papers by U.S. authors was selected from *Physics Abstracts* 1969 in each of 11 subfields of condensed-matter physics. In each of these subfields, a number of experts of trustworthy judgment were identified and asked to grade the scientific impact of each paper with which they were familiar, a number −1, 0, 1, 2, or 3, according to the scale described in Table XIV.12. It was hoped that this scale would have a fairly objective meaning and that, by asking the experts to grade papers in their own specialties, we could avoid any bias for one specialty over another. About two thirds of the papers in the sample received grades, and a sizable minor fraction of these were graded independently by from two to five different experts. Although the agreement of these independent grades was far from perfect, there was some correlation, which was especially strong in the assignment of grade 0 (grade −1 was very rare). Independently of the grades, the number of citations of each paper in 1970 was determined from the *Science Citation Index.*

TABLE XIV.12 Definitions of Grades Used in Study Described in Section 4.7

Label[a]	Description
−1	Work that has an adverse effect on the progress of science.
0	Work with no perceptible impact; that is, the field would have advanced just as far and just as fast if this piece of work had never been published.
1_A	Work that contributes perceptibly to the progress of the field, but whose significance is ephemeral; that is, it is not likely to constitute a significant building block in the corpus of scientific knowledge after several years have elapsed. Inaccurate preliminary results, incorrect but suggestive work, and the like fall into this category.
1_B	Collection of data of reasonably lasting validity but that do not at the moment appear to bear on any new insights.
2	Work of moderately substantial value, that is, it can be expected to constitute a useful building block in the corpus of scientific knowledge for some five to ten years.
3	A major advance, that is, an important new insight or discovery, or an especially definitive measurement or theory on a phenomenon of high interest.

[a] A given publication may have different parts that score differently on the above scale. For this study, the rating was taken to be the highest of all the material contained in it, unless the *net* impact of the paper was −1. If the material in a paper was essentially a duplicate publication of material published elsewhere by the author(s), this fact was ignored in the rating, that is, the paper was evaluated on the basis of what it contained.

It was not expected that number of citations would be a very reliable measure of the merit of any individual paper. A similar study a few years earlier had shown that a paper graded 0 by experts could often have many citations, and a paper graded 2 could often have few. But, as the object of the present experiment was to compare not individual papers but the average quality of work in different subfields, we hoped that averages over some scores of papers would show significant correlations of grades with citations. As Figure XIV.36 shows, there are some modest correlations, though not spectacular ones. Perhaps one can hope, however, that a judicious combination of grading with citation counting will enable one to distinguish fields in which there is an undesirably large proportion of mediocre work from fields in which nearly all the work being done contributes significantly to progress.

It is far from obvious that one can meaningfully compare absolute numbers of citations in different specialties. If one specialty has many people working in it, another only a few, will not a paper in the former

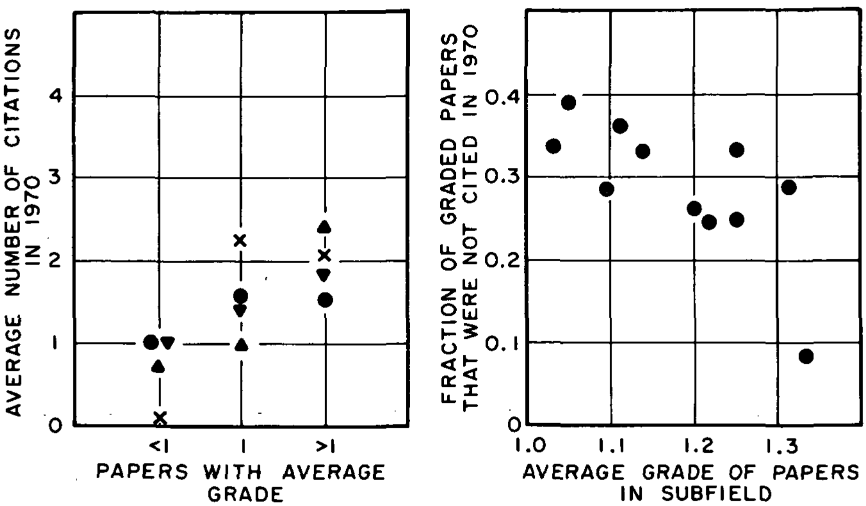


FIGURE XIV.36 Examples of the correlation of quality grades for papers in various subfields of condensed-matter physics with numbers of times these papers are cited. *Top:* the average number of citations for papers in each of three ranges of grades in each of four subfields. (Data for the four subfields in which over 50 papers were graded were used.) Different symbols are used for different subfields. Average grade refers to the average of the grades assigned to a given paper by all the different experts who graded it. *Bottom:* the fraction of papers in each of the 11 subfields of condensed matter that were not cited in 1970, against a rough average grade for these papers.

specialty get more citations than a paper of equal merit in the latter? For specialties small enough that each worker keeps fully in touch with all work that is at all relevant to what he is doing, this situation would undoubtedly be the case. But in larger fields, there seems to be what may be called a "dilution effect": The more publication there is in a field, the less likely it is for a given paper to be cited by any other paper. Thus, in this study, despite differences of more than a factor of five in the numbers of people working in the different subfields, the subfield with the smallest average number of citations per paper (1.20) was a middle-sized subfield, while that with the largest (3.03) was smaller than average; the second largest average number of citations was only 2.11. We shall discuss this dilution effect in more detail in Section 4.8. It will be especially necessary to take this effect into account when comparing quality at different periods of time.

Although only marginally successful at identifying differences in quality level between papers in different branches of condensed-matter physics, the grading experiment gives rather cheering news regarding the health of the subfield as a whole in the United States. It contradicts the subjective view sometimes encountered that "most of the literature is rubbish." Thus, of all the papers graded, less than 8 percent were judged worthless by *any* of the experts grading them (that is, fell in the <1 category of Figure XIV.36), and 33 percent were graded 2 or 3 by one or more of the experts (category >1). From what little data we have been able to gather, we suspect that not all fields of science would fare so well under evaluation by their own experts. And since science is a social phenomenon, growing as new results are accepted by general consensus and refined by criticism,[3] this evidence that leaders in condensed matter respect most of the work done in it is a sign of good health.

4.8 PATTERNS OF USE IN PRIMARY PUBLICATIONS

The principal value of primary publications comes from their being read. This statement is not as trivial as it sounds: An appreciable value of these publications to society comes from their influence on the authors who contribute to them; these authors derive much of their motivation and satisfaction from publishing, and the act of organizing material for publication is surely very useful in clarifying their perception of the problems with which they must wrestle. It has sometimes even been said that journals exist for authors, not for readers. But this is absurd, even though one may grant that journals are not always as reader-oriented as they should be. The interest of authors in publication, through which they receive the benefits just described, would practically vanish if articles in general (not just their own articles) had no readers. The evidence we have presented in Sections 2.1 and 2.2 and Figures XIV.3, XIV.4, XIV.5, and XIV.7 makes an overwhelming case for the importance of primary publications, especially journals, for the day-to-day work of the physicist. And if we are to manage primary publications so as to fulfill this role most efficiently, we must understand what their users do with them.

Let us begin with a brief recapitulation of some of the things we learned in Sections 2.1 and 2.2. In the acquisition of useful items of information, primary publications are by a considerable margin the most important sources both for the information itself (Figures XIV.5 and XIV.7) and for awareness of its existence (Figures XIV.4 and XIV.7); they are also (Figure XIV.7) important as feeders for other final channels. In all these roles, journals loom considerably larger than preprints and reports, but the latter are also important. Journals are useful not only as sources for information that is being consciously sought but also as a browsing medium, in which capacity they are probably, by a narrow margin, the most important source of leads to unexpected information.

One can learn a great deal about the use of primary publications by authors of research papers from studies of citation statistics. [We have already discussed some statistics of this sort in Sections 4.3 (Figures XIV.29 and XIV.30) and 4.5 (Figure XIV.35).] It is of some interest to explore a little further the effect depicted in Figure XIV.35, namely, the way in which citations of papers published in year τ by papers published in a fixed year t decrease with increasing age $(t - \tau)$. As we mentioned there, this decrease is compounded from the fact that the number of papers $n(\tau)$ published in year τ gets smaller as τ recedes into the past and the fact that the older papers become out of date or forgotten. What we would like to investigate now is the function $P(t, \tau)$ that gives the probability that a given paper published in year τ will be cited by a given paper published in year t. (Both papers are assumed to be in some given area of science, whose boundaries are to be considered fixed.) This function is related to the following observable quantities:

1. The average number of references per paper,

$$R(t) = \int_{-\infty}^{t} P(t, \tau)n(\tau)d\tau \tag{4.2}$$

2. The age distribution of references in an average paper,

$$\nu(t, \tau) = P(t, \tau)n(\tau) \tag{4.3}$$

3. The time variation of the total number of citations of an average paper,

$$C(t, \tau) = P(t, \tau)n(t) \tag{4.4}$$

The data we have at hand at the moment are

1. The fact that $R(t)$ seems to have increased by 30 percent or so in the last three decades (see Section 4.5)
2. The fact that, for $t = 1967$, $\nu(t, \tau)$ seems roughly proportional to $\exp(-\tau/4.4 \text{ yr})$ (Figure XIV.35)
3. The fact that, for $\tau = 1962$, $C(t, \tau)$ seems roughly proportional to $\exp(-t/5.7 \text{ yr})$ (Figure XIV.37)
4. The fact that $n(t)$ has been going roughly as $\exp(t/10 \text{ yr})$ (from a guess that its growth rate is about 10 percent less than an average slope in Figure XIV.10, because of widening of the coverage of *Physics Abstracts*).

If we assume the simple exponential forms

$$n = n_0 e^{\alpha t}, P = P_0 e^{-\beta t} e^{\gamma(\tau - t)}, \tag{4.5}$$

the four input data overdetermine α, β, and γ. (The inputs need not be exactly consistent, of course, as they refer to slightly different populations.) The compromise,

$$\alpha = 0.09, \beta = 0.08, \gamma = 0.15, \tag{4.6}$$

fits passably well, however, giving 4.2 years for the time constant in the t

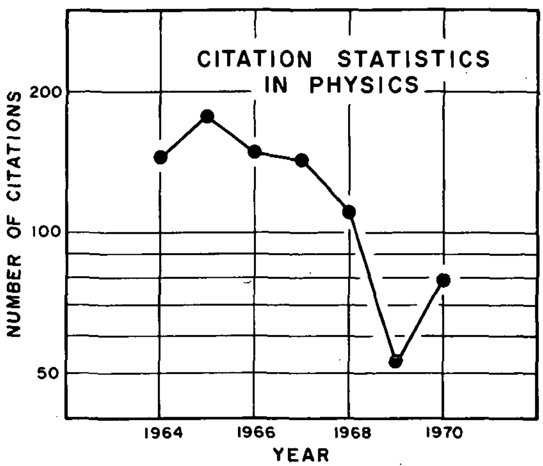


FIGURE XIV.37 Number of citations listed in the *Science Citation Index* in each of seven years for a set of 102 articles published in 1962, half in the *Physical Review* and half in the *Annals of Physics*. (The largest number of citations for any one article in any year was 20.)

dependence of ν (observed ≈ 4.4) and 7.1 years for that in the t dependence of ν (observed ≈ 5.7 years).

The factor $e^{-\beta t}$ [in Eq. (4.5)] describes what we have been calling the "dilution effect": As time advances, all papers, both recent ones and those that have been in existence any given number of years, are less likely to be cited in a particular paper than was the case in former years, because there are so many more papers for them to compete with. This effect almost cancels the rate of growth of the literature, but not quite: β is only slightly less than α.

Figure XIV.37 shows another interesting feature that, though of the order of the random fluctuations in the data, correlates with enough other information to justify being considered real. This is a rise in the citations $C(t,\tau)$ in the first couple of years t after the publication date τ. This increase in citation may represent a diffusion time for general awareness of a paper to develop.

We have given special attention to citations because they are measures of that type of use of the literature that is devoted to the work of incorporating new pieces into the structure of science. But neither this fact nor the relative ease of gathering citation statistics should blind us to the fact that they measure only one type of use. The primary literature may be used by people who are not doing work that is to be published, and even research workers do not cite all the articles they read and benefit from. Statistics on literature consulted in libraries provide one source of information that is relevant to all types of use. Table XIV.13, for example, shows the age distribution of bound volumes of journals left on tables in the Bell Laboratories library: Except for the dip in the 1940's (quite comparable to that for total publication; see Figure XIV.10), the data are

TABLE XIV.13 Age Distribution of Journals Left by Readers on Tables in a Library at Bell Laboratories, 1971

Range of Dates	Observed Number of Physics Journals	Calculated[a] Number
1970–1971	13	0
1965–1969	73	63.2
1960–1964	36	36
1950–1959	32	32.7
1940–1949	4	10.7
<1940	5	5.3

[a] Based on the assumption probability of consultation of a journal from year $\tau \propto \exp(\tau/9\text{yr})$ and fitted to the observed number for the years 1960–1964.

fitted fairly well by proportionality to $\exp(-\tau/9\text{ yr})$. This decay time seems to be slightly longer than for nonphysics journals in the same library; but it is much longer than has been reported elsewhere[85,86] for broader ranges of material.

Comparison of this rate with the slope $n\alpha \exp(t/8.6\text{ yr})$ of Figure XIV.10 leads at first sight to the astounding conclusion that the probability for an article published in year τ to be consulted by a user in year t is practically independent of τ. However, several corrections must be applied that somewhat mitigate this conclusion. As we have noted earlier, the slopes in Figure XIV.10 doubtless overestimate the growth rate of the literature somewhat, as the coverage of the abstract journals of physics has been broadening both in respect to inclusion of semipublished material, the scanning of obscure journals, and the inclusion of work peripheral to physics. Also, the necessity to go to the library to consult a journal is greater if the journal is old, as copies of the more important journals are sometimes available, for perhaps the most recent decade, in departmental and individual collections. This correction factor is undoubtedly large (see following paragraph), but it is at least slightly compensated by the facts that the breadth of the library's collection has been expanding in recent years and that old issues of some journals have been removed to storage elsewhere. Even with these factors taken into account, however, it is hard to escape the conclusion that *the obsolescence half-life of journal articles in physics, as defined by probability of use by any one individual, is probably at least as long as that defined by citations (Figure XIV.35 top) and may well be of the order of a decade.*

The preceding discussion has encountered the question of where the use of published material takes place, an interesting question in its own right. Survey and questioning of a sample of industrial basic-research physicists have indicated that roughly three fifths of their reading of primary publications takes place in their offices and that the remainder is roughly evenly divided between the library and home. While it would be risky to generalize this finding to other populations, it seems likely that reading in a library usually constitutes only a minor fraction of the total time involved. Of course, this figure is not a measure of the importance of library material; in former days, one often borrowed a volume from the library for reading, and today one even more often copies an article to read at leisure elsewhere. Still, it is likely, at least for physicists, that a sizable fraction of the time spent in consulting articles that are not initially available in personal or local-group collections is in fact spent in the library. If this assumption is true, it is obviously very important so to locate libraries and sublibraries, and so to stock the latter, that the already considerable reluctance to use libraries can be counteracted.

One regrettable consequence of the quite understandable tendency of people to rely principally on information sources that are close at hand, rather than journeying to a library, is a provincialism in reading patterns. This provincialism is further enhanced by the reluctance of scientists to read articles in foreign languages, a reluctance that is probably much more marked today than it was a generation ago, when the present predominance of English as a language for scientific publication (see Section 4.3) was much less marked. Some sketchy data, obtained in the course of the study described in connection with Figure XIV.37, suggest that U.S. physicists discriminate against material in French or German by a factor of the order of two or three and against Russian material by a much larger factor. It should be possible, however, to get a much more reliable estimate from well-designed citation studies.

A telling example of the selectivity of current-awareness scanning of journals is provided by records of papers suggested by a sizable population of research physicists for discussion at a seminar of the "journal club" type, that is, one devoted to 10- or 15-minute discussions of items of newly published work that seem particularly novel or exciting. Except for items obtained via preprints or from the organizer of the seminar, who made an especial attempt to spot articles in many journals, a sizable majority of the suggestions originating with other members of the population were found to come from *Physical Review Letters.* There must surely be more work of the type sought originating outside the United States than within it, and as only one fourth to one fifth of the articles in this journal were of foreign origin in the period considered, one would expect that an unbiased scanning of the world's literature for exciting items would get significantly less than half its yield from *Physical Review Letters.* Evidently, the population of physicists involved scanned other journals much less efficiently.

There is still further evidence, though none of it is very quantitative, that many physicists miss a good deal of material of potentially high interest to them because this material is scattered through so many journals that they are unable or unwilling to find it by scanning either the journals or a current-awareness publication. For example, we have spoken with several physicists who habitually scan many journals regularly but who, for some reason, abandoned this scanning for a year or so and then, after another year or two, systematically looked back at the literature of the missing year. In each case, they found many articles of high interest to them that they had not learned about through any other channels such as oral communications, meetings, and cross references in subsequent papers. Although more quantitative studies of missed information are much to be desired, it seems clear that if an effective method of selective dissemination of published information could be devised—conceivably the "user journals" mentioned in Sections 4.3 and 4.9—the average physicist would, for the same investment of time, make significantly better contact with new work of value to him than he does now.

Studies of the distribution of journal articles with respect to number of physicists reading them, skimming them, and reading their abstracts would be very illuminating. Interesting results on such topics have been reported in chemistry[87] and psychology.[88]

Little information is available on patterns of use of preprints and similar material; we have already presented, in Figures XIV.3, XIV.4, XIV.5, and XIV.7 some fragments of information about their relative position in the overall communication picture. These findings suggest that physicists, at least in large institutions, spend an amount of time reading preprints comparable to that spent on journals but that they find the preprints rather less rewarding. The study of high-energy theorists by Libbey and Zaltman[18] contains further data on correspondence between authors and recipients of preprints.

As for the use of reports, as distinguished from preprints, the few studies that have been made indicate that they are cited far less than journal articles but that their use is increasing. Thus, Burton and Green[89] found that in 1959 some 3 percent of the references in the *Physical Review* were to reports, as were 1.7 percent of those in the *Journal of Applied Physics* and 0.6 percent of those in two British journals. In 1971, our counts have given 4.3 percent in the *Physical Review;* in 1936, the percentage was much smaller. Anthony *et al.*[38] have added further observations.

4.9 SUMMARY AND OUTLOOK

There has been so much material to discuss on primary publications that by now the reader may well feel both weary and confused. So it is appropriate to recapitulate the most important facts and reflect on their implications for users of information, for research organizations, for societies and publishers, and for sponsors of research.

New research is presented in a wide spectrum of journals, large and small, prominent and obscure, general and specialized (see Figures XIV.17, XIV.18, XIV.27, and XIV.28). A small fraction of research—but much more than a generation ago—is published in conference proceedings and other books. Specialization of journals has increased greatly in the last few decades (Figure XIV.28), but many journals covering all of physics continue. Large journals published by societies have a more dominant position in physics than in most other sciences, especially in the United States, and are growing more rapidly than the literature as a whole (Figure XIV.20). Many important journals are also issued by commercial publishers, especially in Europe, as are, however, many minor and highly specialized journals. Circulation, which has a significant inverse correlation with price per kiloword (Figure XIV.19), is largest for the U.S. society journals, which can be marketed very inexpensively because of the support of much of their prerun costs through page charges. Such a system of sharing of production costs between the sponsors of research and the users of journals gives society a much better return on its total investment than pricing at a level sufficient to recover both prerun and runoff costs (see Section 4.2). Improved technology has largely offset rising factors in production costs in recent years, and the circulations of the society journals continue to be high. However, some vigilance, perhaps including the so-called "two-track" system (Section 4.2) is needed to ensure maintenance of page-charge income (Figure XIV.23) in times of decreased funding. This should be possible without significant hardship to anyone, given an understanding approach by publishers, authors, and those who support research.

Though some commercial publishers have supplied valuable and needed services through their alertness to needs for specific types of journals, others have started unnecessary journals that are immensely expensive for the amount of material they contain (off the scale in Figure XIV.24) and are a regrettable financial drain on library budgets. Is it too much to hope that authors can be educated to avoid journals whose prices are out of proportion to the benefits they offer? After all, the authors themselves benefit from the larger circulations provided by less expensive journals of high quality.

Patterns of publication have been changing slowly in the last few decades. Papers are becoming longer (Figure XIV.33); the number of papers per physicist per year has been increasing (see Section 4.1), as has the number of authors per paper (Figure XIV.34). There is a moderate and probably increasing publication of papers originating in one country or region in journals produced in another (Table XIV.9). English has come to be the predominant language of publication throughout most of the world, and the most important sources of non-English publication—Russian journals—are in many cases available after a few months in English translations. In typical subfields of physics that have been studied, the quality of work currently published in the United States seems to be high, in the sense that only a very small fraction of the publications fail to have an appreciable positive impact on the leaders who set the pace of the field (see Section 4.8). The rate of obsolescence of physics papers, as judged by studies of citations (Figure XIV.35) or of journal use (Table XIV.13 and discussion in Section 4.8), is slower than has often been supposed.

Despite all these favorable facts, there is evidence that primary publication is not fulfilling its role in communication as well as it should. There are a distressingly large number of journals, and their availability in personal collections is rapidly decreasing. Though reprography has been a boon for study of papers once they have been identified, it is becoming more difficult for the physicist to maintain awareness of all the work that might be of high interest to him (see Section 4.8); the "dilution effect" described in Sections 4.7 and 4.8 implies that even his probability of picking up subsequent references to this work in other papers is decreasing.

Various measures have been proposed to counteract these trends. One is large-scale subsidized distribution of preprints to open-membership "exchange groups"; this has been hotly debated[82,83] and proposed again in a more cautious form.[18] As argued in Section 4.6, we feel that no quantitative expansion of the present scale of distribution is desirable but that some systematization of a limited-distribution system might be worthwhile, although it would never be the solution for the overall problems of primary publication. The purveyors of secondary services have pinned their hopes on more extensive and rapid current-awareness publications, wide availability of computer tapes, and improved subject indexing. Although measures of this sort offer a possibility of improved access to the journal literature, their effectiveness will depend greatly on the scale and diligence of their use, and it is difficult to extrapolate from the present low use of secondary services (Figure XIV.3) to a reliable prediction of how improved services will be used. A third suggestion[74] is the repackaging of papers being published in conventional journals into "user journals" oriented toward any of a large number of specialties and issued nearly simultaneously with the original journals. This scheme, which we have discussed (Section 4.3), would yield a product sufficiently similar to that of conventional journals that user response should be quite predictable; the problems are economic and political.

5 Advanced Books, Reviews, and Compilations

In this chapter and in Chapter 7, we shall consider the consolidation of research-front knowledge for more effective use by research workers and advanced students. We shall defer until Chapter 8 the discussion of how this knowledge is reworked for use in other scientific disciplines and in technology as well as by lower-level students and the public. These stages of reworking can be most effectively carried out only if they have been preceded by a consolidation of the primary research results, so it is appropriate to consider the latter first. We shall start by explaining why such consolidation is becoming increasingly necessary if research workers (and *a fortiori* anyone else) are to be able to make efficient use of all the new knowledge that is being generated. Sections 5.2 through 5.5 will describe the present situation: the general amount and character of the various kinds of written syntheses of physics information that are now being produced (5.2); their intellectual coverage, adequacy, and currency (5.3); the processes by which they are created (5.4); and their distribution and use (5.5). In Section 5.6, we shall discuss some characteristics of this literature and its use that should be goals of its production, and in Section 5.7 we shall suggest possible measures to achieve these goals.

(As much of the large-scale work of data compilation is done in or for information-analysis centers, most of the discussions of this aspect of consolidation of information will be deferred until Chapter 7.)

5.1 THE NEED FOR CONSOLIDATION

The sifting, evaluation, and compaction of new knowledge have always been a central part of the advance of science.[3] In the "old days"—say, before 1940—many of the greatest minds in physics contributed enormously to the science by gathering together, in one mighty treatise, essentially all that was known about some area and presenting it from a unified and critical point of view. One has only to think of Maxwell's *Treatise on Electricity and Magnetism,* Lamb's *Hydrodynamics,* Sommerfeld's *Atombau und Spektrallinien,* the review or Handbuch articles of Sommerfeld and Bethe (Elektronentheorie der Metalle, in the *Handbuch der Physik*), Bethe, Livingston, and Bacher (Nuclear Physics, in *Reviews of Modern Physics*), and the like. The important role of such works, some of which are still useful today, has been eloquently described by Goudsmit.[90] Unfortunately, in the present generation the tremendous expansion of the literature of physics has made it almost impossible to put together such complete syntheses. But even more unfortunately, it has greatly increased the need for them, or indeed, for any competent consolidation of information.

The way that times have changed is illustrated schematically in Figure XIV.38,[91] a figure that reflects the experience of physicists who have been working for several decades and that is partially confirmed by such facts as the dilution effect (see Section 4.8). The full curves represent the level of awareness of a typical research worker, Mr. A, in various fields x, plotted as a function of the intellectual remoteness of x from A's specialty. Because of the vastly greater rate of production of new information now than

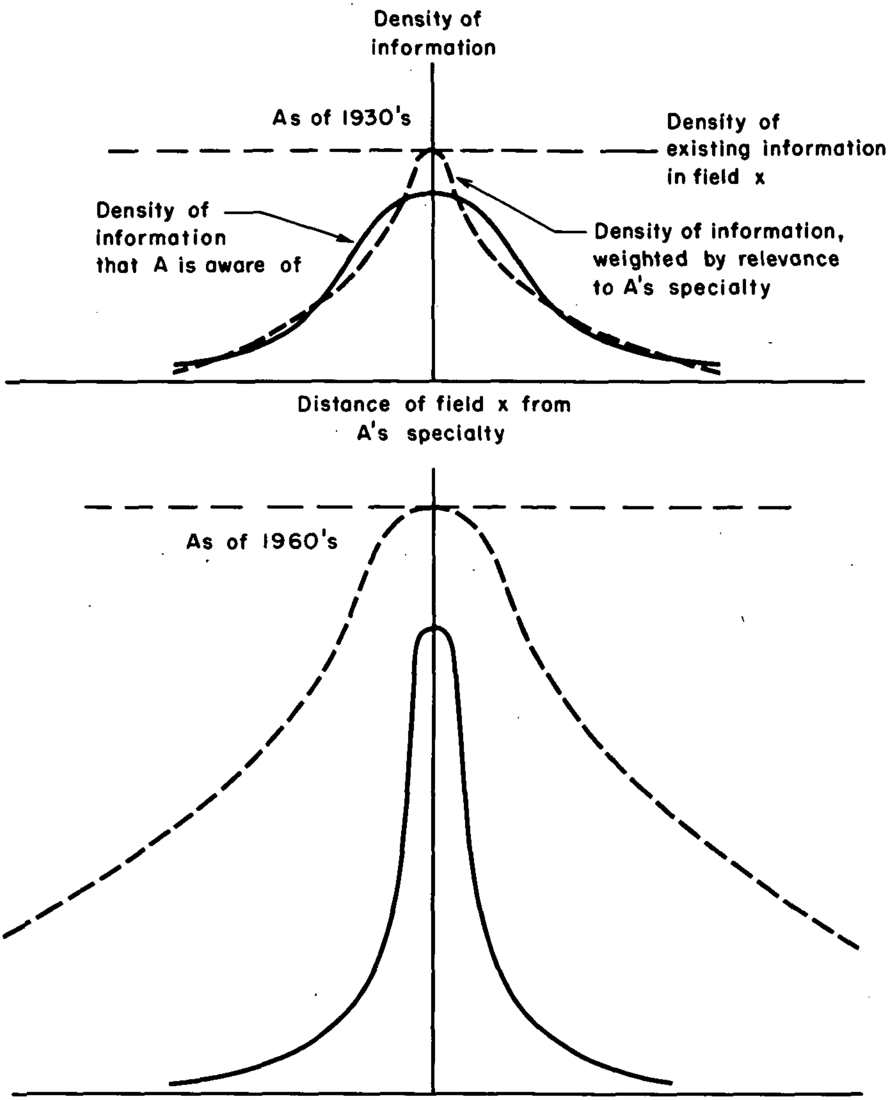


FIGURE XIV.38 Schematic diagram depicting awareness versus relevance for fields neighboring the specialty of a typical physicist, Mr. A.

in the 1930's, the lower full curve has to be much narrower than the upper, the area under it being little if any greater. The dashed curves, on the other hand, represent the amount of information in different fields that is reasonably relevant to A's specialty. Advancing knowledge is continually establishing relations between fields that were previously unrelated. So the dashed curve for the 1960's is broader than that for the 1930's. The full and dashed curves have changed in opposite directions, and their relation to each other is now, in most fields, qualitatively different from what it was then. And the situation will surely be worse in the future.

The seriousness of the difficulty depicted in Figure XIV.38 varies enormously from one individual to another; it is probably worst for theorists trying to fit concepts and observed facts together and is less serious, though usually still troublesome, for many experimentalists and for theorists who only compute. It would be most interesting if a rough quantitative measure could be obtained for the area between the full and dashed curves in the lower part of the figure, and, though difficult, it does not seem impossible to design experiments to yield this information. At present, only fragmentary data are available. For example, one can get a lower bound, doubtless many times too small, from the few missed-information studies already alluded to in Section 1.1 (none of them in physics). More significant, probably, are some of the examples given in the latter part of Section 4.8, which illustrate that a systematic combing of the journals will produce many items of high interest to a typical physicist that he does not learn about in his normal pattern of activities (if, as is usual, this normal pattern does not allow time for systematically searching all the journals or an equivalent current-awareness service).

The evidence of Figures XIV.4, XIV.5, and XIV.7 indicates that awareness of current information of value comes in comparable measure from scanning the literature, which most physicists take time to do only incompletely, and from personal connections—mostly preprints and oral conversations. We have just cited evidence that the latter, valuable though they are, do not afford complete coverage of available information. However, the future outlook for the role of interpersonal contacts is probably even worse for two reasons. One is that as more interrelationships among fields develop, the wings of the dashed curve of Figure XIV.38 continue to grow in importance. But exchanges with colleagues are most frequent and most productive in relation to areas that are near one's central area of interest; they are less so in the wings. The other consideration is that for economic reasons the exponential growth of information we have been experiencing probably will slow down considerably in the next generation. As this deceleration proceeds, there will be an increase in the proportion of the total store of information, relevant to a given piece of current work, that is more than a few years old. This older information is less likely, as compared with fresh information, to emerge from casual contacts with one's colleagues.

Evidence of the sort described in the preceding paragraphs seems to indicate, then, that in many fields the information channels used at present—including the extremely important informal person-to-person contacts—are not going to suffice to cope with the growing flood of information. As we have seen in Section 3.6, much thought is currently being given to improved retrieval mechanisms and to selective dissemination. Both of these are essential if we are to make any sort of progress. But when the wings of the dashed curve in Figure XIV.38 are very broad, only a highly sophisticated retrieval or selective-dissemination system could suffice to keep a specialist in a given subfield in touch with all information in neighboring subfields that is relevant for him.

But this is not the most serious problem. It will often happen that even a perfect supplying of all relevant documents will be of little use. Herring[91] has mentioned an extreme example, having to do with the theory of the effects of orbital quantization on electronic transport in semiconductors. Even within this narrow specialty, there are between 100 and 150 papers in the literature, most of them of great mathematical complexity. Some are wrong, some are trivial, but most do contribute something of value. All that is of value in all of them could be presented much more lucidly in one tenth of the space, yet no small group of these papers could be selected that would cover all the ground. Rarely, if ever, will even an expert theorist undertake to digest all this material; the interested experimentalist can only take the time to read at most a very few of the papers; therefore, he cannot get any sort of perspective on what is known. For such a case, even the optimum retrieval or selective-dissemination system would merely provide the user with the 100–150 papers; it would not and could not ensure that the real information content was perceived and assimilated.

Although this example may be extreme, something of the sort must happen in a great many areas of physics. According to the evidence presented in Section 4.7, the great majority of papers published in physics contain something of value and should not be ignored by those whose interests they touch. A common case is that of data evaluation: A research worker or technologist who needs a particular numerical datum will be poorly served by being given a bibliography of some dozen apparently conflicting measurements of the quantity in question. In some areas of physics, the need for systematic and critical data compilations is becoming serious. One is nuclear physics. A recent study by a panel of the National Research Council[92] has pointed out that compilation of nuclear data has fallen dangerously far behind the accumulation of experimental results.

A final point to be made on the general problem of the need for information services is that one must always distinguish between the information *needed* by a scientific worker (that is, all information that could be of positive help to him) and the information consciously *wanted* or *sought* by him[93]; both, of course, are to be distinguished from the information actually *obtained*.

To summarize then, it should be clear that the functions that reviews and compilations are supposed to perform—digesting, evaluating, consolidating, simplifying, and repackaging for specific classes of users—are essential if information is to be effectively used. Suitably designed studies could probably put an order-of-magnitude dollar value on an ideal review literature (as defined in Section 5.6) for any given field. In some areas—for example, undergraduate teaching—supply-and-demand economics already gives us an idea of the value and provides reasonably efficient economic incentives for production of the material. In other areas, notably at the research frontier, there are too many impeding factors to allow supply-and-demand optimization to work. The present situation is probably so far from the ideal that it is metastable. In this situation lies at least a partial justification for the fact that the present section shows what may seem to be a rather one-sided preoccupation with research workers and their needs; however, this area seems to be the one in which the needs are most poorly perceived and the problem of filling them most difficult.

5.2 AMOUNT AND TYPES OF SYNTHESES NOW BEING PRODUCED

One of the most obvious questions one can ask about the existing review literature is: How much of it is there? However, this question is not as simple as it sounds. It needs to be subdivided, and some pitfalls in simple ways of answering it need to be avoided.

What we have been calling "review literature" can be subdivided into review articles, treatises, and compilations. Although the boundaries separating these categories are a little fuzzy, most items fall fairly clearly into one or another of them. However, the gathering of data on the amounts of material of each of these types is difficult, since review articles appear in review journals, research journals, and hard-cover books (such as conference and summer school proceedings), compilations often occur in unpublished reports, and there is a continuous gradation from reviews and treatises that are useful to the research worker to those that are only for popular or student use. The rough and incomplete data presented in this section will refer to material useful to research workers and, where possible, will be subdivided according to the distinctions just mentioned.

A study made for the National Science Foundation a few years ago[94] gathered statistics on the numbers of review articles in selected fields of science and technology, largely by counting articles identifiable as reviews in the abstract journals of these fields. These and other raw data need to be processed by taking account of the selection criteria of the different abstract journals, their overlap with each other, and like factors. For example, the ratio of number of review articles to number of primary publications appeared to be about 0.012 in physics and 0.061 in chemistry, as determined by counts in *Physics Abstracts* and *Chemical Abstracts,* respectively. But, whereas most of the listings in *Physics Abstracts* are on physics, rather than a borderline field, the situation in *Chemical Abstracts* is the reverse, and the excursion into borderline fields is wider for reviews than for primary papers. Rough counts[90] (Herring 1966, unpublished data) of individual papers combined with an even cruder allowance for the imperfect coverage of review articles in hard-cover books by *Physics Abstracts,* suggest that in core areas of physics and chemistry the ratios may have been more like 0.023 and 0.037, respectively.

There is evidence that ratios of this order apply in most of the pure sciences except mathematics, where the review article—as opposed to the treatise—is little used. But as one goes further toward everyday applications—as in technology and clinical medicine—one finds a higher proportion of reviews, presumably because of the many different types of repackaging of existing knowledge that are needed by different groups of users. We must remember this fact when we discuss, in Chapter 8, the communication of physics information to engineers and others.

The difficulty in making reliable estimates even from so clear-cut a source as the subject index of *Physics Abstracts* is illustrated in Figure XIV.39. The fraction of the total entries that were classified as "reviews" increased in the years immediately following the introduction of this subject heading in 1969 and has since been decreasing. In what proportions is this latter trend due to a decrease in review relative to primary writing, to expansion of coverage of nonreview material, to ever poorer coverage of review material, and to changes in the definition of "review" for indexing? That the last factor might well be serious is indicated by two independent samplings of *Physics Abstracts* entries, the one (1969 abstracts) used for Tables XIV.10a, XIV.10b, and XIV.11 of Section 4.5, the other (1967 abstracts) done at the American Institute of Physics.[95] Both studies found the proportion of review articles among the entries to be approximately twice as high as the corresponding points of Figure XIV.39; there is some evidence that the ratio is now closer to unity.

Of more interest than numbers of review articles are the total number of pages written in reviews, treatises, and compilations, the distribution in size of the articles or books, and the further characteristics we shall discuss in Section 5.3. Figure XIV.40 shows how the picture for physics compares with that for a few other fields. The data were obtained from counts in ab-

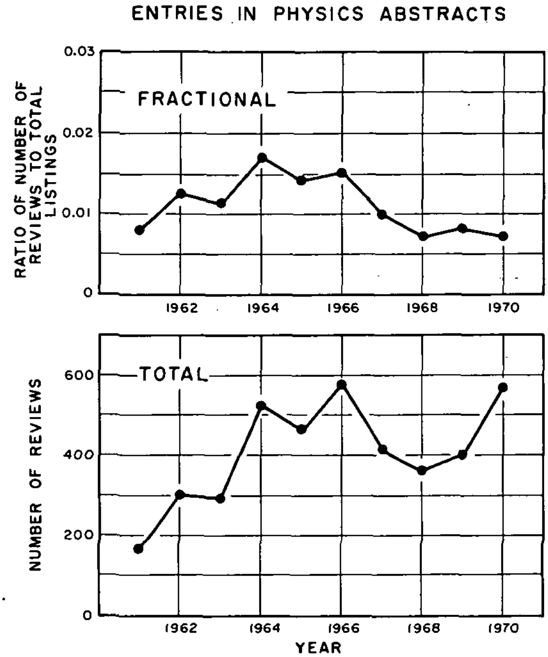


FIGURE XIV.39 Total number (lower graph) and fractional number (upper graph) of entries in *Physics Abstracts* that were indexed as reviews for each year from 1961 to 1970. (See text for discussion of the significance of the numbers.)

stract journals and counts of "books received" in *Nature, Physics Today, Science,* and other similar publications[91]; the estimates are extremely crude, but an effort has been made to eliminate overlaps, so that the totals in the various boxes can be legitimately added. A more detailed picture of the relative amounts of different kinds of review literature is given in Figure XIV.41; the two halves of the figure differ both in the methodology of estimation and in the areas covered (all physics versus solid-state physics). How much of the difference between them is due to the one factor and how much to the other is not quite certain; at any rate, the similarities and differences between the two charts give one a feel for the extent to which conclusions from them can be trusted. It is especially difficult to decide where to stop counting books as they get more elementary or more peripheral. Undergraduate textbooks have been included in Figure XIV.40 but not in Figure XIV.41. A perhaps extreme example of the difficulty with peripheral books is shown in Figure XIV.42; the peripheral material is bigger than the field.

A few points emerge from Figures XIV.40, XIV.41, and XIV.42:

1. Not only is much more consolidated knowledge published in books (above the undergraduate level) than in journals, but there is even a greater

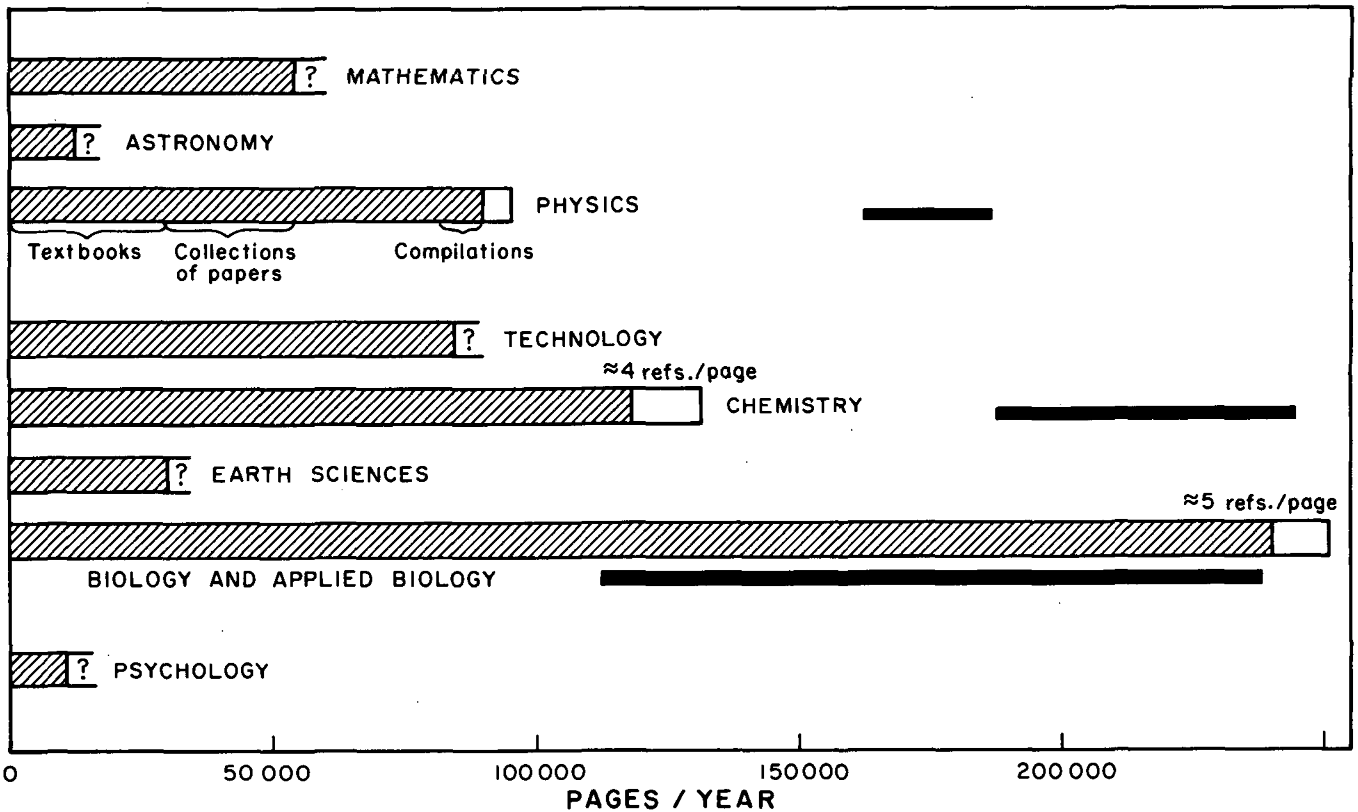


FIGURE XIV.40 Rough estimates[91] of the annual page bulk of review literature in several fields, as of 1966. Open bars are review articles in journals, shaded bars, material in books of various kinds. Solid bars give estimates of 10 percent of the primary literature. Material pertaining to border areas between two fields has been assigned to only one of them, so the numbers for different fields can be added without overcounting. A few rough figures are also given for the average ratios of bibliography entries to pages.

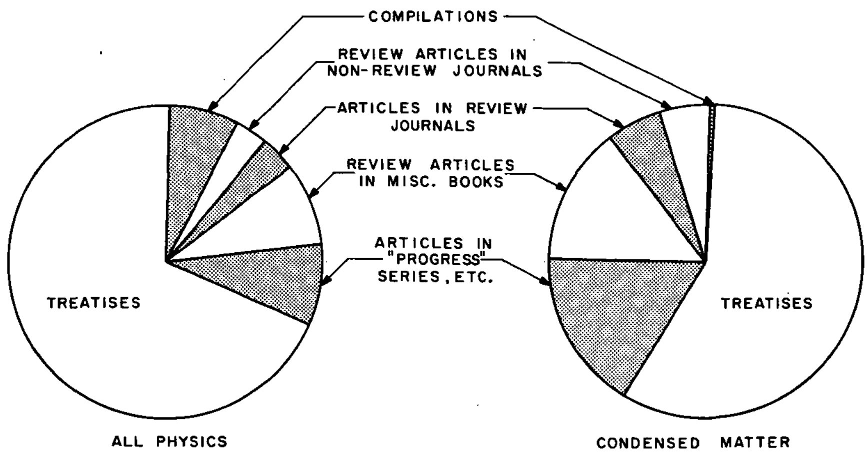


FIGURE XIV.41 Two kinds of estimates of the relative page bulk of different kinds of review literature, in physics as a whole and in solid-state physics. *Left:* Estimates for all of physics, 1966,[96] obtained from analysis of *Physics Abstracts,* book lists, and the like. *Right:* Estimates for solid-state physics, 1946–1967, obtained from querying a sample of solid-state physicists in their role as authors.

bulk of review articles in books than in journals. These are in series such as "Progress in . . ." or "Annual Review of . . .", in conference proceedings, summer-school lectures, and the like, or occasionally in other types of multiauthor books.

2. Review articles in journals are distributed about equally (in total pages) between review journals and journals that publish mainly new research. (As the review articles in the latter journals tend to be shorter, they are actually more numerous.)

3. In solid-state physics, and probably in the rest of physics, too, the total bulk of pages in reviews, compilations, and treatises of use to research workers is roughly one fifth that of the primary literature.

4. The number of pages of review-article literature that are required to survey a given number of research papers is greater in physics than in most other fields of science, as might be expected from the highly conceptual nature of physics. Treatises use more pages per research article cited than do review articles, perhaps because review articles often do not give details of the arguments of the papers they review, and perhaps because treatises often give a personal, rather than a synthetic, development of their subject.

No accurate information is available on the language distribution of the material just described, but it is fairly certain that English predominates

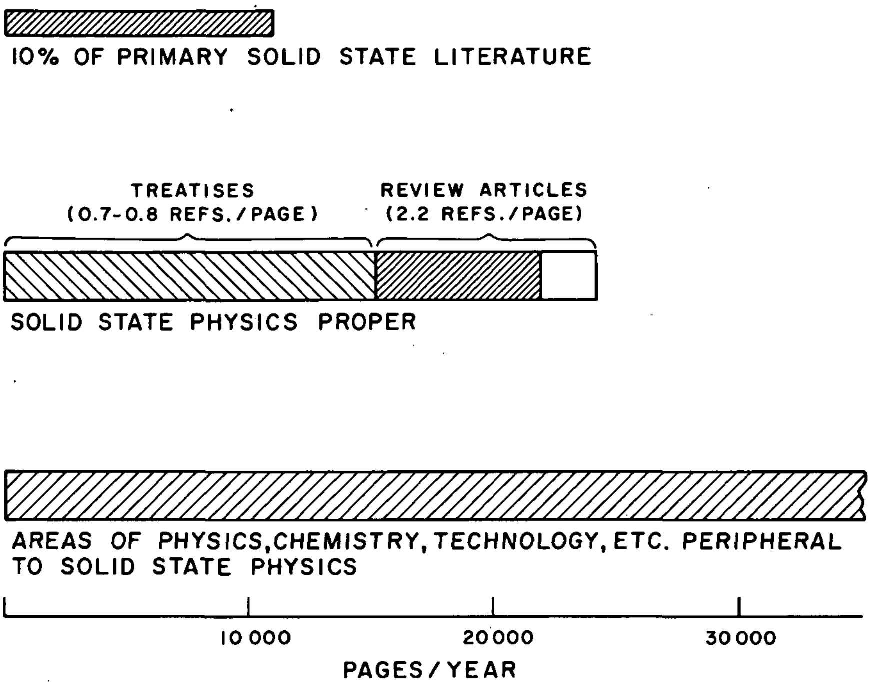


FIGURE XIV.42 Bulk of various kinds of review material relevant to solid-state physics produced in 1966, as estimated from analysis of *Physics Abstracts*, book lists, and the like. As in Figure XIV.40, the shaded broad bars represent material in books, the open bar, review articles in journals. The narrow bar at the top shows 10 percent of the primary literature in the field. (Note the general consistency with the independent estimate on the right of Figure XIV.41.)

to about the same extent as in the primary literature (Section 4.3). The largest and strongest single review journal in physics is the Soviet *Uspekhi Fizicheskikh Nauk*; its review articles are reproduced in English translations by the American Institute of Physics. A comparison of book listings in *Nature* with listings in *Uspekhi* suggests that, as of 1965, there were perhaps 15 percent as many physics books published in Russian as in English.

5.3. COVERAGE, STYLE, AND OBSOLESCENCE

The obvious next question is, how much of the journal literature do these treatises and review articles cover? We shall discuss two variants of this question: First, how many journal articles are referenced? Second, how complete and how deep is the coverage of *significant* ideas in the literature?

It is easy, though tedious, to count the number of papers cited in any book or review article. Figures XIV.40 and XIV.42 have already given some rough average values of the number of different articles cited per page of review for different disciplines and different types of review material. Table XIV.14 gives some more detailed statistics for physics. We note:

1. The density of citations is several times higher in review articles than in treatises; it varies significantly from discipline to discipline, being higher in more empirical or taxonomic fields and lower in the more theoretical or logically structured ones.

TABLE XIV.14 Sample Figures on the Average Density of Citations in Different Types of Review Literature in Physics

Category	Sample	Items or Articles in Sample	Approx. Average No. of References in a Kilo-word of Material[a]
Treatises	Solid-state treatises listed in *Physics Today*, Sept.–Oct. 1966	10	1.7
	Miscellaneous sample[b]	4	1.2
Review articles in books of the "progress" type	*Magnetism and Magnetic Materials Digest* 1966	20	11.7
	Progress in Elementary Particle and Cosmic Ray Physics Vol. 8	4	4.7
	Annual Review of Nuclear Science (part of Vol. 16)	6	7.5
	Solid-State Physics. Advances in Research and Applications	12	3.9
	Magnetism (Rado and Suhl, Eds.)	36	3.2
Review articles in other books	Solid-state reviews in conference proceedings listed in *Physics Today*, 1966	≈15	4.4
Review articles in journals	Sample from *Reviews of Modern Physics*	5	3.8
	Sample of review articles from *Physics Abstracts* (Stern[95])	43	3.7
	Miscellaneous sample[b]	8	4.9
Compilations	Sections of physics interest in *Landolt-Börnstein*	≈150	5.0
	Nuclear Data Tables 1960	9	4.7
	Miscellaneous sample[b]	1	5.7

[a] Based on the assumption of ½ kiloword per page in most books, 1.0 in *Reviews of Modern Physics* and *Landolt-Börnstein*, and 0.75 for the average of journals from the *Physics Abstracts* sample.

[b] A small sample of review material of all types pertaining to magnetism and semiconductors, 1967–1968.

2. Even in a subfield like solid-state physics, where the proportional bulk of the review literature and the number of citations per page are both fairly low, the total number of citations in a year's crop of books and reviews is half again as large as the total number of original papers published in the same field in a year. However, this figure refers merely to the sum of the numbers of entries in the bibliographies of these books and reviews. Not all of the papers in these bibliographies are in the field selected (solid-state physics), and the different bibliographies doubtless have a great deal of overlap.

Much more important, and at the same time more difficult to answer, is the second question: How well are the really significant ideas and findings covered? To attack this problem, one needs a criterion of significance and a scale of coverage. The only study seems to be one in the field of solid-state physics, briefly mentioned by Herring.[91] In this study, he investigated the coverage in treatises and reviews of some samples of papers whose significance was high enough to put them in the top tenth or so of all papers published. Although a number of levels of coverage of these papers were studied, it will suffice here to quote the data pertaining to a "passing level" of coverage. The coverage of a particular journal article in a review is said to be of "passing level" if, at the very least, the major results of the journal article are quoted. (Almost two thirds of all solid-state reviews citing a given journal article discuss it at this "passing level.") From extensive sampling of both books and review articles, Herring concluded that the probability for a typical one of the significant papers to be discussed at the "passing level" in a treatise or review within five years after its appearance is in the range 0.5 to 0.9, a value of the order of 0.8 being most likely. Although this finding seems at first sight encouragingly high, it appears to result only from an unpredictable diversity of coverage in very many items of review literature; the most respected and comprehensive single source, the multivolume series *Solid-State Physics: Advances in Research and Application*, has only a probability of the order of 1:6 to cite a given significant item at the "passing level" within five years.

Similar information about the coverage of the review literature in other fields would be most useful. Because of the technical judgments required on questions of significance and adequacy of coverage, any attempt to dig out such information will require the active participation of experts in the field being investigated.

Treatises by their nature try to develop a moderately broad area in depth; they are aimed, in comparable numbers, at graduate students, research workers in a field, or scientists in other closely related fields who need an understandable account of the given field for their work. Only a minority of treatises attempt comprehensive coverage of the literature. Most acknowledge that their treatment is selective; often only one of several schools of work is discussed. Review articles are even more diverse. Those in books of the type "Progress in . . ." or "Annual Review of . . ." usually undertake fairly comprehensive reviews of the literature. The median length of these articles, as judged from the samples used for Figure XIV.41 (described more fully in Section 5) and Table XIV.14, is of the order of 40 book pages, with the more prestigious series and multiauthor treatises running rather longer. Review articles in journals range all the way from very brief surveys of two or three pages, usually for the nonspecialist, to long comprehensive articles of the sort appearing in books of the "Progress in . . ." type. The latter are more likely to be found in review journals. From the samples just mentioned, it appears that the median length for articles in review journals is of the order of 25 pages of *Reviews of Modern Physics* (equivalent to about 50 pages in a modern book) and of eight or ten pages for reviews in nonreview journals.

A review or book is never quite up to date even at the moment the author completes it, as it takes time to work new results into the treatment. The time lag in publication sets it even farther behind. Both types of delay are apt to be worse for treatises than for review articles. Figure XIV.43 shows one way of measuring their combined effect. Although the data are rather sketchy, they show the expected behavior; the references in articles in a review journal have a peak frequency at an age of about two years, while those in books seem to peak at a rather greater age. (If Figure XIV.42 were extended farther into the past, all the curves would surely go down with a slope intermediate between that for growth of the literature—a little less than the slope of Figure XIV.10—and the steeper slope resulting when this is compounded with obsolescence, as in the lower curves of Figure XIV.35.) The multiauthor book is apt to be the slowest of all, as the completed manuscripts of some authors sometimes have to wait for the contributions of other authors.

But what about the subsequent obsolescence of books and reviews? Although this is related to the obsolescence of the primary literature that the reviewer attempts to summarize (a topic we have already discussed in Sections 4.5 and 4.8), it is not quite synonymous with it. A good book or review may retain its value for many years because of its profound perspective, completeness, or, especially, clarity of exposition of basic concepts. On the other hand, even a quite satisfactory one may almost disappear from use very quickly after another one appears that is better in every respect. As with the primary literature, citation studies give a useful perspective.

Table XIV.15 gives some very sketchy figures on the ages of books and reviews cited in a subsample of solid-state papers taken from Volume 154

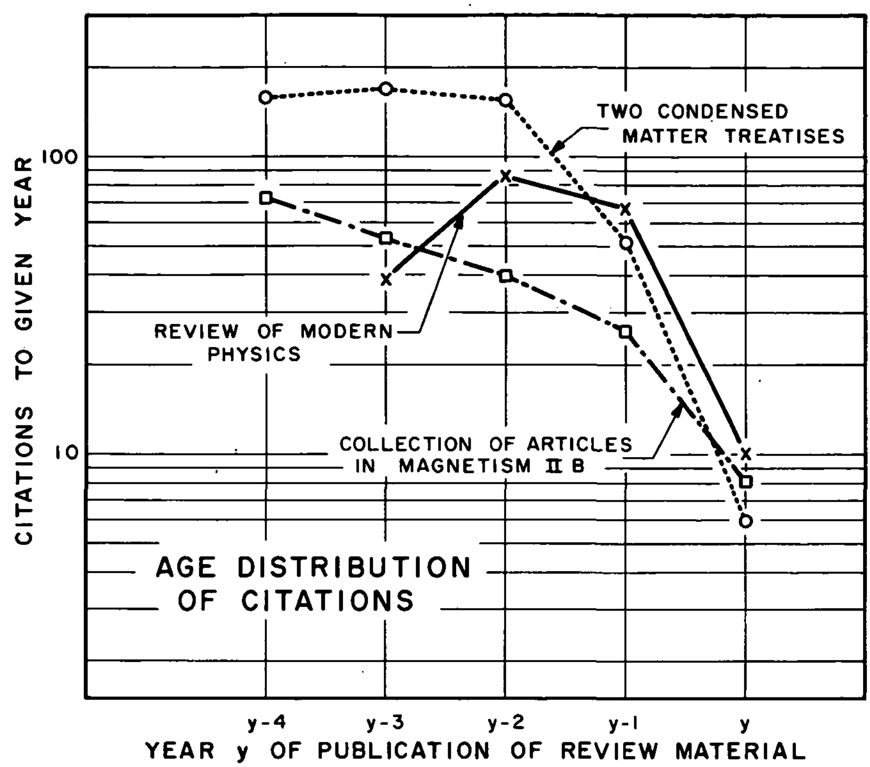


FIGURE XIV.43 Age distribution of citations in the bibliographies of three samples of review literature. Full curve: sampling of review articles from *Reviews of Modern Physics* 1964–1967. Dot-dash curve: sampling from the collection of articles in *Magnetism II B.*[97] Dotted curve: Sampling of references in two solid-state treatises.

of the *Physical Review*. The data were too meager to reveal any possible difference in longevity between books and reviews. At first glance, the data suggest a half-life of about five years. But it must be remembered (see the analogous calculation in Figure XIV.35) that the rate of publication of books and reviews, like everything else, has been growing. If we assume this growth in the solid-state field to have had about a ten-year doubling time, as does physics research literature in general, the half-life for obsolescence goes up to ten years. Although this quick calculation of the correction figure is not accurate, it appears that the half-life is surprisingly long. A possible reason may be that the review literature is so inadequate that old references must be used for lack of anything better.

The picture of a surprisingly long half-life for review literature agrees

TABLE XIV.15 Distribution of Ages of Solid-State Books and Reviews Cited in a *Physical Review* Sample, 1967

Age of Book or Review Cited (= 1967 – date of publication)	Number of Cases
1 to 5 years	20
6 to 10 years	11
11 to 20 years	7
>20 years	2

with counts of citations of articles from *Reviews of Modern Physics* by papers in the computerized TIP file (supplied to us by the American Institute of Physics) after removal of citations to nonreview articles. It is further crudely confirmed by some counts of citations of particular review articles (from the Seitz-Turnbull *Solid-State Physics* series, from *Reviews of Modern Physics*, and from *Uspekhi Fizicheskikh Nauk*) in the *Science Citation Index* 1965. The citations per review article were not perceptibly lower for review articles published from 1956 to 1958 than for those published in the same place in 1964. The contrast of this behavior with that shown for research papers in Figure XIV.37 testifies to the slower obsolescence of the review articles.

5.4 AUTHORS AND THE WRITING PROCESS

Another class of questions concerns the authorship of reviews. Who are the people who write each of the various types of review literature? How much time do they spend on it? How do they go about the task and what are their problems and their rewards?

According to some counts we have made for several samples of reviews and treatises in physics, over half the American books and longer review articles have as an author a Fellow of the American Physical Society (a rank to which about one tenth of the membership has been elected) at the time of publication of the book or review. As the Fellows thus typify the scientists most qualified by talent and experience to write reviews providing some perspective over a sizable area, a natural experiment[96] was to query a random sample of Fellows (in solid-state physics) to find out how many and what kind of reviews they had written. The results suggested that the men in this group had been writing review material (including advanced textbooks) at a rate of the order of 15 pages per year per man. This figure is roughly consistent with the data of Figure XIV.40 and the estimate that approximately half the review writing done by U.S. physicists is done by Fellows. Again, Hagstrom[31] has obtained data from a sample of academic physicists. He found theoreticians and experimentalists wrote an average of 0.28 and 0.10 review article a year, respectively, while the authorship of books in their careers to date was 0.52 and 0.32, respectively. These figures are consistent with those of the Herring sample.

The investment of time required for the preparation of treatises, reviews, and compilations, of course, varies greatly from case to case, even if expressed in terms of hours per page of final product. All the authors questioned agreed that the ratio of hours to pages is considerably lower for review writing than for the production of a research paper (research time plus writing time). However, for a critical review the difference in time may not be great. A figure near this end of the scale is provided by the experience of one author in the preparation of two very long reviews; he expended about seven hours for each printed page (500 words) of the final product. A more typical figure would probably be less than this. Thus, in the writing of books and reviews for use of other research workers, the average expenditure of time by solid-state physicists of the level of APS Fellows may be roughly 50 or 60 hours per year.

These statistics having suggested that review writing is not in general at a level that cuts seriously into the time of most of the competent physicists, it was natural that more detailed studies of the motivations and problems of authors of such material should be undertaken. Several such studies were conducted in 1970–1971 by the American Institute of Physics and the Subcommittee on Reviews and Compilations of the Advisory Committee on the AIP Information Program. Among the findings of these studies were:

1. For review articles, sampled from *Physics Abstracts,* the median duration of active work on the article was of the order of three months; another six months or so typically elapsed before appearance in print.
2. U.S. authors of a representative distribution of books, articles, and compilations seemed to receive some personal compensation (honoraria or royalties) in about half the cases, but the amounts were practically always minuscule in relation to the time devoted to writing.
3. Although nearly half the authors in the sample just mentioned acknowledged some governmental or other outside support for their work, the preparation of the review, book, or compilation was rarely explicitly identified among the objectives for which the support was granted.
4. Motivation for undertaking reviews, books, and compilations seemed to come in comparable degree from the authors' own perceptions of need and from external urging (for example, invitations from editors) or the

occasion of giving a lecture or course; in about one third of the cases, both factors were simultaneously important.

5. In only a minor fraction of the cases were graduate students or postdoctoral associates enlisted to help with the work; yet nearly two thirds of the authors said that they would have used such assistance if funds to pay for it had been available.

6. A sizable minority of the respondents mentioned that it would have been advantageous to have additional bibliographic or editorial assistance or funds for computer time.

Points 3, 5, and 6, especially, have been further supported by an independent study of the Division of Nuclear Physics of the American Physical Society. This study revealed a widespread willingness of the leading workers in the field to supervise the preparation of nuclear data compilations if adequate postdoctoral assistance could be provided. The outcome has been a program that we shall describe briefly in Section 5.7.[92]

5.5 DISTRIBUTION AND USE

Remembering that the amount of use made of any information resource is strongly correlated with its accessibility, we ought now to look at the circulations of books, review journals, and other consolidations and at their market prices, which undoubtedly influence circulations.

Let us look first at books. Figure XIV.44 gives some information on prices and prices per page for various kinds of English-language books containing consolidated information of interest to research workers in physics (of course, not all of the material in these books is of this type—for example, conference reports contain many primary research papers). We note:

1. Prices of books vary widely from below 1¢ to over 8.5¢ per page.
2. The median price is around 4.5¢ per page, or 9¢ per kiloword.

Information on circulations of books is usually not available. Our impression, based on information supplied by a handful of authors, is that books of the types shown in Figure XIV.44 typically have total sales in the range of 2000–3000. A very successful graduate text may run considerably higher; an outstandingly successful "Progress in . . ." series may reach 5000. The estimate of U.S. expenditure on physics books used in Section 2.3 was obtained as follows: We assumed half the total circulation of the books of Figure XIV.44 to be in the United States, and a curve of average circulation against price going from 3600 at $2.50 to 1600 at $65. Inte-

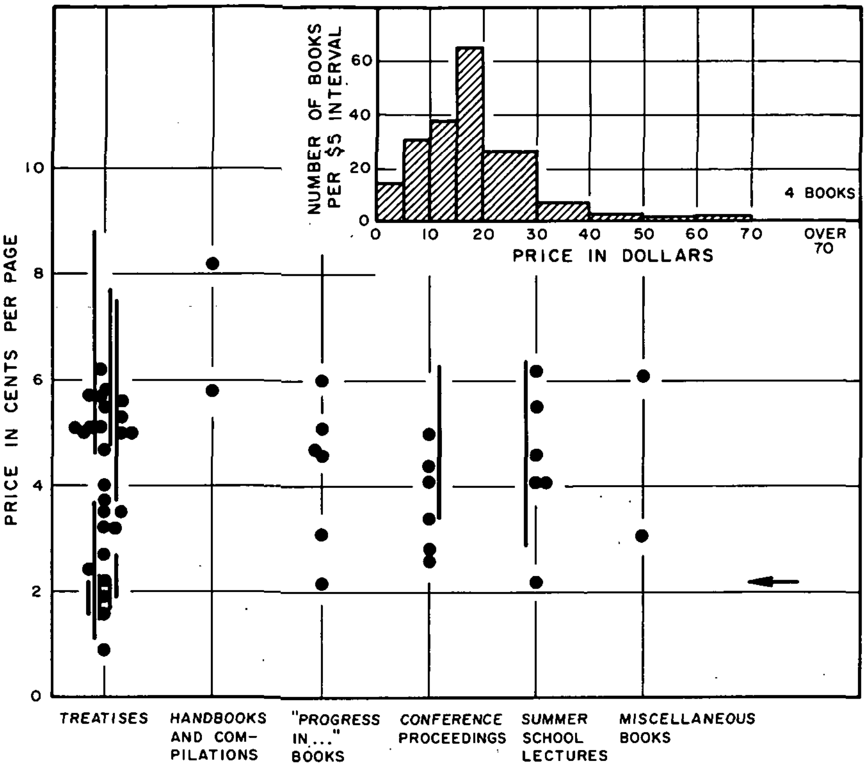


FIGURE XIV.44 Scatter diagram of prices per page for a random sample of physics books of various types listed in *Physics Today*, 1970. When two prices were quoted (hard cover and paperback), these are represented by the ends of a vertical line segment. For most books, a page is about half a kiloword, so these prices should be doubled for comparison with Figures XIV.19 and XIV.44. The arrow at the right represents the median price of a sample of eight elementary textbooks. Insert at top: histogram of price distribution of all books listed in half of the last year that were considered to be in physics proper and of interest to research workers.

gration over the histogram at the top of Figure XIV.44, with additions for foreign and unpriced books, gave a total purchase expenditure of $11 million.

Data on review journals are still more meager, as there are few of them and their circulations are not always available. Figure XIV.45 gives some price and circulation figures for a few of them. We note:

3. Some commercially produced review journals have prices in the range 9–10¢ per kiloword, very similar to those of books.

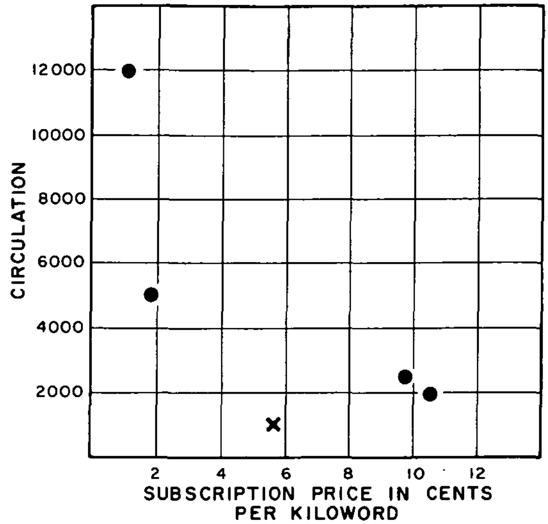


FIGURE XIV.45 Prices and circulations of a few review journals in physics (as of 1968 or 1969). The highest point is *Reviews of Modern Physics;* the next highest is the Russian *Uspekhi Fizicheskikh Nauk;* the X point represents the translation journal, *Soviet Physics Uspekhi.*

4. Circulations of these review journals are rather larger than for primary journals of similar unit price (Figure XIV.19).

5. Subsidized review journals, sold at a low price, can have vastly larger circulations.

The low circulation of the translation journal *Soviet Physics Uspekhi* (the X point in Figure XIV.45) follows the pattern found for translation journals generally (Figure XIV.19). Yet the quality and price of the material offered would seem to justify a higher circulation. Perhaps, despite the many good physicists in the Soviet Union, there is an impedance mismatch between Soviet and Western schools of thought; perhaps the Western physics community and its libraries have merely been slow to appreciate this source of review articles, despite its having been available for over a decade. Citation statistics, interestingly, show an exaggerated preference for *Reviews of Modern Physics* by U.S. authors and for *Uspekhi* by Russian ones.

The general picture one gets from these fragments of data is that, although *Reviews of Modern Physics* is available to most research physicists in their own offices, their neighbors' offices, or in some nearby subcollection, comparable availability exists for no other English-language review journal, and for few, if any, books. Subsidized publication, enabling sale at approximately runoff cost, is probably necessary if a review publication is to be this widely distributed. The social desirability of such subsidy is obvious. The same types of arguments that were used in Section 4.2 apply also to the review literature.

Now let us turn to the actual use of books, reviews, and compilations. We have seen in Section 2.1 and Figures XIV.3, XIV.4, and XIV.7 that physicists (beyond the graduate-student level) seem to spend only a small fraction of their time, of the order of an hour a week, in the use of these sources of information and that only a small proportion of the useful information physicists assimilate in their daily work comes from them. Certainly, we are still far from the day envisaged by Bernal[93] and others, when users of information will get nearly all of it from treatises and reviews without seeing the primary publications.

What are the reasons for this smaller role of books, reviews, and compilations relative to primary literature and interpersonal contacts? Certainly, an important reason has to do with the related factors of time lag and coverage. We have seen in Section 5.3 that a sizable minority of the items of significant information, five years or so old, have still not been covered at the "passing level" (quotation at least of major results) in the consolidation literature. Moreover, according to Figure XIV.43, even a book that has just come off the press will usually be quite deficient in its coverage of the literature of the last two or three years, and a fresh review may be deficient for the last one or two years. Referring to Figure XIV.8, we see that something like half of all useful items of information currently used by physicists are no more than a year or so old at the time they are apprehended. Thus, for a sizable majority of the items of useful information that physicists currently acquire, books and reviews *could not* at present be made to substitute for the other sources used. But it still may be true that they are underutilized. Figure XIV.8 suggests that they could play a much larger role than they seem to in the transfer of information that is already fairly old at the time it is acquired by the user. And if their effectiveness in this role were greater, much more old information would be acquired and used.

This reasoning suggests that when we look at citations of review literature, we do so with an eye to finding out what types of such literature are most used. One type of data that can be obtained is illustrated by Figure XIV.46,[96] which shows the distribution of citations in the *Physical Review* over the various kinds of review literature and contrasts this distribution with that previously presented in Figure XIV.41 for the bulk of review literature produced in the various categories. We note:

6. Treatises account for a large proportion of the citations but do not dominate the citations as they do the figures on total review pages written. This may be because the treatises contributing to the latter are not oriented so completely toward the research worker as the review articles are, and in part also because the treatises take longer to publish, so are less up to date when they appear.

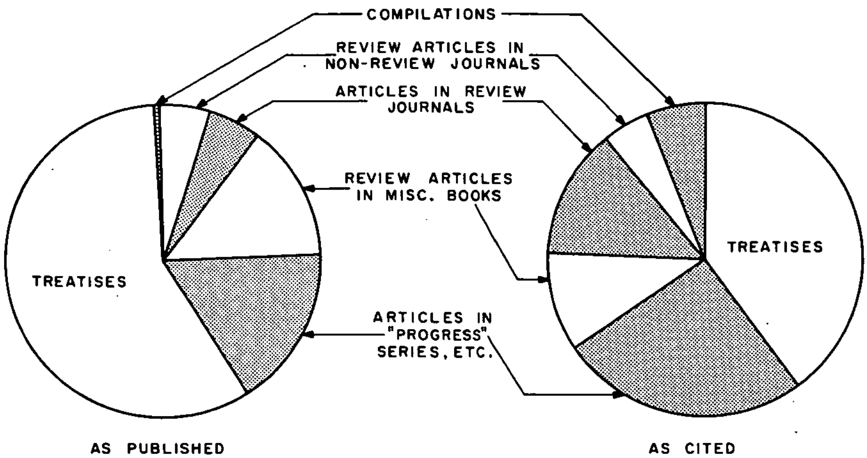


FIGURE XIV.46 Comparison of the relative page bulk of different kinds of review literature in solid-state physics (1956) with the relative numbers of citations to these same kinds in part A of the *Physical Review,* 1964.

7. Equally noteworthy is the fact that the review articles appearing in nonreview journals and in conference reports, summer school notes, and the like seem to be less often cited in proportion to their bulk than the longer ones appearing in "Progress in . . ." series or review journals.

These two points are confirmed by two further studies.[96] One consisted in asking a number of research physicists the question, "What books or reviews have you found especially useful as sources of information about the last decade of developments in your field?" Of 61 items named, 22 were treatises, 20 were articles in books of the "Progress in . . ." type or in the *Handbuch der Physik,* 13 were review articles in review journals, 2 were review articles in conference or summer school proceedings, 2 were review articles in nonreview journals, and 2 were compilations. The other study, based on the sample of review authors mentioned at the start of Section 5.4, gave the results shown in Table XIV.16. As will be seen, points 6 and 7 are again evident. There seems to be a general and not unexpected tendency for the number of citations of an item of review material to increase with its length more slowly than proportionally. More important, however, and this we feel is the principal import of point 7 and Table XIV.16, *the review articles that are the most used are ones that have reasonable thoroughness and are published in fairly obvious and accessible places.* An author serves himself and his colleagues poorly if, after making the effort to write an acceptable review, he publishes it in an out-of-the-way book

TABLE XIV.16 Citations, in *Science Citation Index* 1956 and 1966, to Books and Reviews Written by a Sample of U.S. Solid-State Physicists

Category	Number of Items (Books or Articles)	Ratio of Citations to Total Pages
Treatises	9	0.052
Articles in books of the "Progress in . . ." type or in multiauthor treatises	20	0.13
Articles in conference and summer school proceedings	21	0.08
Articles in other types of books	4	0
Articles in review journals	10	0.46
Review articles in nonreview journals	15	0.05

or in a journal one does not normally think of as a repository for review material.

5.6 SOME GOALS

Let us now turn from these statistics on what we now have to a consideration of what we would like to have. Desirable characteristics for review literature are, of course, easy to enumerate: It should be easy to locate, comprehensive in coverage, authoritative, objective, critical, well organized, lucidly written, up to date, and so on. What is needed is not so much to enumerate these qualities as to assess the importance of each of them compared with the effort required to secure it. The following paragraphs will offer a few considerations that should enter into such assessments. They will be largely qualitative, though some of them could and should be made roughly quantitative by suitably designed operational-research experiments. Although studies of user preferences (as, for example, the survey of chemists by Brunning[98]) can provide valuable hints for the design of such experiments, they are, in many respects, too unreliable to be used alone in the formulation of goals for review literature.

5.6.1 *Accessibility*

For a book or review to be useful to a user, he must know of its existence and be able to obtain it. Even today, awareness and access are far from easy to achieve. Most of the review literature one has occasion to use pertains to the wings of the relevance curve (dashed) in Figure XIV.38. It thus falls in areas in which the density of the user's current awareness (full curve) is low. As the review literature is published in many journals and hardcover collections as well as in single treatises, it is often difficult to locate.

Abstract journals, citation indexes, and book listings, though all sometimes useful, are not really satisfactory for this task. It would be much better if one could turn to a standard place listing all books and reviews in one's field, arranged according to subfields with appropriate cross-indexing, and stating for each item the nature of the audience for which it was written, the date of the production, the length, and the approximate extent of the material covered.

5.6.2 *Coverage*

There are disadvantages as well as advantages to exhaustive coverage of the literature of the field being reviewed. In many areas, it is impossible to find a single author with the time and patience to review all the papers in a specialty definable by a title of the usual length. Unless the topic to be covered is further restricted by qualifying language, anywhere from a sentence to a paragraph in length, the literature on it may run to thousands of papers. In many fields it has, therefore, become the custom for the author of a review to weed the literature of his topic down to a fraction of its size before starting to digest it. This practice not only saves work for the reviewer but makes the review more concise and easier to assimilate.

Sometimes this sifting and evaluating is done systematically and conscientiously, but too often the selection is determined by the reviewer's more superficial tastes, his friendships, and his laziness; often the full bibliography is not even identified. As long as such cases exist, the user can never be sure whether a paper not mentioned by the reviewer is wrong, is of minor importance, or is a significant development that the reviewer overlooked. If such doubts strongly assail the user, he must delve repeatedly into the primary literature, and a sizable part of the utility of the review is lost. In an ideal review literature, there would be a class of reviews with exhaustive coverage of an area precisely delineated at the start but so organized as to devote a minimum of space to the discussion of papers of little value, thus allowing the real "meat" to be assimilated very quickly.

Incidentally, there is often a real value to the discussion of wrong or foolish papers, beyond the value of warning people not to read them. For if other researchers can be educated not to make the same mistake, their time will be saved.

5.6.3 *Need for Different Levels of Reviews*

Besides the exhaustive reviews just mentioned, there is a great need for reviews covering a broader area in less detail. This need, which has been widely recognized, arises in part from the requirement of the research worker to sift the information of value to him in areas a little removed from his main specialty (the wings of the curves in Figure XIV.38) and in part from the desirability of providing scientists with a general education in fields neighboring their own. Thus we need about three levels of reviews. It is usually assumed that reviews at these three levels will be written and published quite independently. However, it may be worthwhile to give more thought to several possible relationships among them.

It would be helpful to the user if each treatise or review article identified itself at the outset as belonging to one of the three levels just mentioned and gave further information describing the audience to which it was addressed. Book prefaces usually do this now, of course, though often not concisely; review articles often do not. A review not of the exhaustive type but of the second level in detail of coverage should provide an exhaustive referencing of the available exhaustive reviews in the areas it deals with. With such material at hand, the research worker could quickly spot areas outside his own specialty that might contain material of value to him and could delve into them at whatever depth was necessary.

Entirely different books or articles are not always necessary to provide different levels of reviews. It may be possible, for example, for the author of an exhaustive review to include clearly identified summarizing sections that can be read independently of the remainder of the material and so serve the purpose of a review at the next level. When it is feasible, this type of organization can not only save manpower in the preparation of reviews but can also prove a convenience to the user who is interested in the exhaustive review.

5.6.4 *Organization of Review Material*

A review of any great length will often be used piecemeal. Users outside the specialty to which the review is devoted will usually be interested only in a part of the material covered, unless the review is of the sketchy type intended for general education of nonspecialists. Users who are specialists in the field of the review will already be familiar with most of the material covered and will not want to take the time to read through a lot of familiar material to find the few things that will be new to them. Only the worker who is just starting to specialize in the field of the review will wish to read every word, and even he will so wish only if the review is more authoritative than others in the same field. It is, therefore, of overriding importance that treatises and longer review articles be so organized that small portions of them relevant to any particular interest can be quickly identified and then assimilated with a minimum of backtracking.

It is easy to identify several traits that facilitate such piecemeal use. One has already been mentioned: summaries of different portions of the review, so arranged that they can be read and understood independently of the main text. Another is the clear identification of a place, or at most a few places, where the definitions of mathematical symbols and the like are given. Still another is the relegation of auxiliary data and arguments to appendixes.

The treatment of the bibliography deserves special attention. Often the user consulting a treatise or a review article of the longer type merely wants to find out how the reviewer evaluates or explains a particular research paper. It is often hard to locate the discussion of a particular paper, unless an author index is provided, and even then it is sometimes not very convenient. An ideal system is to have an alphabetical bibliography at the end of the book or article, in which the title of each paper cited is given, together with a list of the page or pages of the review on which it is discussed.

5.7 SOME SUGGESTIONS

5.7.1 *Speed of Publication*

While the long half-life of reviews with respect to obsolescence, which we inferred in Section 5.3, might seem to suggest that prompt publication is unimportant, one must remember that in physics, most of the previously unknown information one uses is of rather recent origin (see Figure XIV.8). Thus, as we have surmised in Section 5.5, the relatively low present use of books and reviews is probably due largely to their not being sufficiently current. Improvements in the speed of publication of review material would, therefore, improve its utility significantly.

5.7.2 *Aids and Incentives to Authors*

How can one get qualified experts to write more and better reviews? This is obviously the biggest question. It is usually, and doubtless rightly, assumed that the problem is to improve the review output of those people who, through their current activity on the research front, are in the closest contact with expanding areas of knowledge. Before taking up this problem, however, we should direct some attention to possibilities for the use of people who specialize in review preparation, rather than in research. One such possibility occurs for some types of compilations, when there exists an organization big enough to maintain a large continuing effort in the area of compilations (for example, the U.S. Bureau of Standards). Again, it has been suggested[90] that the actual writing of reviews be done by specialists in writing, who use material and judgments supplied by one or more research experts; these writers must, of course, have enough background in the field to understand the points the expert wishes to convey. They, like the compilers, would have to be employed at a center of some sort, supported by a scientific society or government agency, though perhaps often located at a university. (We shall return to this question in Chapter 7.) Finally, there is the suggestion that senior scientists or technologists no longer active in research might well devote their energies to review writing. In areas where innovations have not been too rapid, and occasionally in other areas, this plan might work and would have the advantage of giving the reviews a longer and often broader perspective than they would have if prepared by younger authors.

The reluctance of active research workers to undertake the preparation of reviews and their predilection for skimpy rather than thorough ones arise mainly from the laboriousness of the task. As an example, Way[99] has given a colorful description of the difficulties besetting the author of a data compilation. Anything that will reduce the labor will improve the quantity and quality of the product. A sizable fraction—though usually a minor fraction, at least in physics—of the work involved in review preparation consists of locating and arranging references, numerical calculations, preparation of graphs, and other rather mechanical tasks. The author might be significantly aided if these could be performed by the staff of a center, maintained by a scientific society or other organization, or by suitable financing of a staff at his own institution.

Along with the laboriousness, an important practical deterrent to the preparation of reviews and compilations is the fact that, at present, they rarely have a place in the responsibilities for which scientists and their assistants are explicitly supported (see Section 5.4). *It should be as easy for a scientist to obtain grants for review preparation as for new research, and such grants should allow for participation, when appropriate, of graduate students and postdoctorals, and for travel, computer time, overhead, and the like.* A specific proposal for an experimental program of grants of just this type has recently been made by the American Institute of Physics[100]; even farther advanced is a two-year program, administered by the National Research Council, with National Science Foundation support, for compilation of nuclear data by teams of senior physicists and postdoctoral fellows.[92]

Of greater potential, though of uncertain practicality, is the possibility of using, in a single review, the expertise of a large number of authorities. This procedure is, in fact, almost essential if a review of the exhaustive type is to be prepared covering more than a narrow area. Usually, such reviews are prepared as multiauthor treatises in which the subject is brutally chopped into subdivisions, each of which is attacked independently by one of the participating experts. The disadvantages of this practice are obvious: There is no unity of viewpoint, and often part of the literature of the subject is missed because none of the authors considers it in the little province assigned to him. Perhaps thought should be given to ways of making it possible for one expert to write a comprehensive review, but with many other experts participating in the collection and critical evaluation of the material. Suitable ways of compensating these experts, financially and with recognition, would have to be devised.

5.7.3 *Aids and Education for Users*

We have suggested in Section 5.5 that potential users of books and reviews are often unaware of the existence of those that best fit their needs. Therefore, we would like to reiterate the recommendation in Section 5.6 that systematic preparation of annotated bibliographies of reviews be undertaken. In addition, a more systematic attempt to teach graduate students how to use the review literature of their field—and indeed, to use all bibliographic tools—might be helpful.

Finally, we should note again the social desirability of making review literature of all types available to prospective users at as close to runoff cost as possible (see Section 5.5). To achieve this goal, ways should be found to subsidize not only its authorship but, in many cases, its publication.

6 Oral and Other Interpersonal Communication

We now turn to the important lower part of Figure XIV.1, which, according to Figures XIV.3 through XIV.8, plays a role in the communication of physics information comparable in importance to that of the journal literature. This part comprises all oral communication, which is rather clearly divided into informal (person-to-person) and formal (lectures, meetings) channels, and also written personal communication (letters). Although the importance of these media has been widely recognized, little quantitative study has been made of them, in the physics community, at least, because data are not so easy to obtain as, for example, for journal literature. Nevertheless, we shall see that by assembling the previously available facts and augmenting them with a few new ones, a reasonably intelligible picture emerges.

We shall start in Section 6.1 with a look at the most studied aspect of formal oral communication, the scientific meeting. Section 6.2 will take up the other type of formal oral communication, lectures or seminars. In Section 6.3, we shall pass into the rather nebulous area of person-to-person communications, both intraorganizational and between people from different organizations or different countries; in either case, the communications can be either face to face or by means of the telephone. Finally, in Section 6.4, we shall discuss written informal communication.

6.1 MEETINGS

As every scientist knows, there is a fantastic diversity of meetings. Most of those that have to do with physics are listed in the calendars published in *Physics Today;* a more comprehensive announcement service covering all scientific fields can be purchased from the World Meetings Information Center (Chestnut Hill, Massachusetts). To get a little perspective, we should start with a look at the members and sizes of the different kinds of meetings (Table XIV.17). From the table, we learn:

1. Although "miscellaneous" meetings, especially the aperiodic ones, are far more numerous, the regular periodic meetings of scientific societies probably attain nearly half the overall attendance ("man-meetings"), because of their large size.
2. Similarly, the periodic international meetings, especially the ones sponsored by the various Commissions of the International Union of Pure and Applied Physics, probably account for over half of the man-meetings involving foreign travel by U.S. physicists.
3. Total attendance by U.S. physicists at domestic meetings probably amounted to approximately 40,000 to 50,000 man-meetings in 1969.
4. Total attendance by U.S. physicists at conferences abroad is more difficult to estimate but was probably within ±50 percent of 5000 man-meetings.

Although the local expenses of meetings are usually sufficient to require subsidy from registration fees, society funds, government grants, or donations by patrons, by far the largest item of cost (to society as a whole) of meetings is the travel expense for the participants. This varies widely, depending on the distance traveled and the length of stay at the meeting. In computing the figures we used in Figure XIV.9 of Section 2.3, we have arbitrarily taken an average figure of $150 per man-meeting (including participants who do not have to leave home) for domestic meetings. The integrated cost of attendance at foreign meetings we have estimated at about one third that at domestic meetings. Combination of these estimates with the numbers obtained from Table XIV.17 gives the order-of-magnitude number $8 million to $10 million for the annual investment of U.S. physics in meetings, as of 1969.

The sizes of society meetings seem to grow in time about as society membership. Figure XIV.47 shows some examples. The figure also suggests the plausible inference that the proportion of society members who attend

TABLE XIV.17 Approximate Numbers and Characteristics of Scientific Meetings in the United States in 1969 of Interest to Sizable Numbers of Physicists[a]

Type	Number, 1969	Typical Size	Estimated U.S. Physicist Man-Meeting Total for Year (thousands)	Examples
Periodic				
International[b]	5	500–2000	4	International Conference on Thin Films; International Conference on the Physics of Electronic and Atomic Collisions
National meetings of AIP member societies	32	250–1000 (up to 6000 for APS)	20[c]	Society meetings and meetings of APS Divisions
National meetings of other societies	18	–	1	American Association of Physicists in Medicine; American Nuclear Society; Solar Energy Society
Regional society meetings	15	100–400	3	APS Sections; N.Y. Academy of Science
Periodic topical conferences	5	250–800	1.5	Magnetism and Magnetic Materials; Gaseous Electronics
Gordon Research Conferences	8	120	0.7	Chemistry and Physics of Solids
Aperiodic				
International[b]	2	200–1000	1.5	High Energy Reactions
Domestic	61	50–400	8	Laser Safety Conference; Metal-Ammonia Solutions; Hypernuclear Physics; Shock Tubes; Regge Poles

[a] From listings in *Physics Today*, supplemented by a few other sources.
[b] Sponsored by an international organization or jointly by several organizations in different countries.
[c] Based on exact figures from most of the societies.

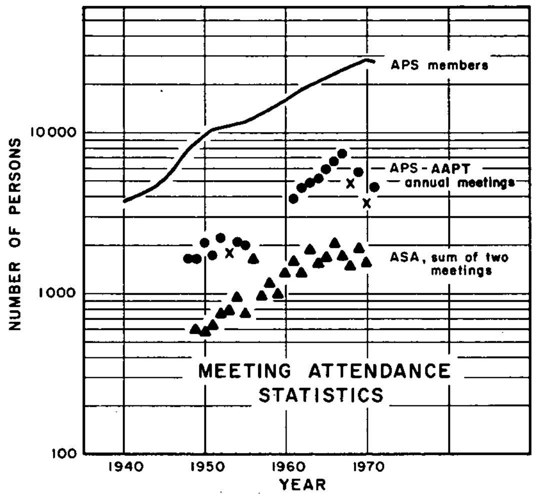


FIGURE XIV.47 Comparison of attendance at meetings with society membership over the last two or three decades. Upper curve: membership of the American Physical Society. Circles: registered attendance at the Annual Meeting of the APS and the American Association of Physics Teachers in New York (unregistered attendance is a small but unknown fraction of this total). Crosses: registered attendance for years when the Annual Meeting was held in a city other than New York. Triangles: sum of registered attendance at the two yearly meetings of the Acoustical Society of America (membership 4403 in 1968).

a meeting rises when funds are plentiful and drops in times of retrenchment. (However, as the data refer only to official registrants, one wonders what proportion of the drop in 1969 and 1970 may have been due to people who attended but felt unable to pay the registration fee.) There has been no noticeable tendency for society members to shun meetings as they get larger, despite complaints that they are "too big." For both the large American Physical Society and the much smaller Acoustical Society of America, the largest meeting in the middle 1960's drew about one fourth of the membership.

What is the role of meetings in scientific and technical communication? This question has been the subject of an extensive research program at the Johns Hopkins University Center for Research in Scientific Communication. We have cited earlier some of the activities of this program dealing with written communication; their studies of meetings were modeled on earlier studies of the American Psychological Association[28,29] and were carried out for several scientific and technological societies, including one in physics, the Optical Society of America.[101] For the comparison of the different fields, see Compton[102] (brief) or Garvey[103] (more detailed). For the Optical Society, the principal meeting studied was a rather large one, with 2248 registrants and 234 papers. Studies were made by questionnaires (with about 314 returns) of speakers, listeners, and persons subsequently requesting copies of papers, and of the interactions of these groups with each other.

The study just described contains a good deal of useful though not surprising data on the educational level of attendees and speakers, the extent to which results reported had been described previously in seminars or reports, the number of review papers, complaints about the size and pattern of the meeting, and the like. One or two items of this sort are of special interest to us:

1. Over three fourths of the authors of papers presented had *specific* plans for formal publication of their material, and half the remainder anticipated eventual publication.
2. For the three fourths of the authors just mentioned, submission of their papers for publication was planned on the average for some time in the three months or so immediately following the meeting (sometimes it preceded the meeting).

Other findings of particular relevance to our chapter are those on the nature, extent, and value of the information communicated at the meeting or subsequently growing out of it. This communication can be divided into that associated with the presentation of papers and that due to miscellaneous interpersonal contacts. Concerning the former, it was found:

3. Over one third of the authors of the papers presented had some interaction, due to their presentation, that led them to modify some of their current activities (most often the conduct or planning of research).
4. About one seventh of the authors made contact, as a result of their presentations, with new people with whom they planned to continue to communicate.
5. Written versions of talks, or related information, were requested from most of the authors (1 to 51 requests, median 8). These requests came before, during, and (most frequently) after the meeting.
6. About one fourth of the listeners at any particular session planned to contact at least one of the authors reporting at that session to get further information on his work, and some 7 percent planned continuing interaction.
7. Nearly half of those attending modified some of their work activities (most often methodology or instrumentation) as a consequence of papers, symposia, or other formal activities of the meeting. Invited and contributed papers contributed about equally to these modifications, though only 15 percent of the papers were invited; exhibits, symposia, and the like contributed a little also.

As for effects of personal contacts not having to do with papers presented on the program, it was found:

8. About one quarter of those attending, that is, slightly less than in 7 above, reported modification of their work activities as a consequence of informal contacts at the meeting.

Although it may be a little risky to extrapolate from these findings in the single field of optics to meetings in all other areas of physics, some general conclusions are at least suggested by the findings. In Figure XIV.7 the role of talks at meetings in the total supply of useful information, though perceptible, could not be accurately gauged because of the rather large statistical fluctuations; in the subjective judgments of Figure XIV.5, it seemed fairly favorable. Points 5 through 7 above, especially 7, give semiquantitative support to these judgments. But it is fairly clear that listening to formal talks is far from the whole story. Points 3 and 4 show that authors receive valuable reactions. And even at this Optical Society meeting, which was very poorly set up for informal contacts, these had over half as much effect on nonauthors as the papers. In better-designed meetings, it would be easy to believe that the informal contacts could be more important. An assessment of the total long-term effect of meetings would have to assign them a significant role in feeding, via the interpersonal contacts initiated and renewed at meetings, the parts of Figure XIV.4 having to do with personal correspondence, sending of preprints, and the like. Thus the total contribution of meetings to communication is undoubtedly considerable. However, we must remember that the benefits of meetings accrue mainly to those who attend them, and these are a rather select group. For example, the Johns Hopkins studies show them to be rather more highly trained and more research-oriented than the average members of the society. Alternatively, we may note that for the American Physical Society, for instance, the average member goes to 0.40 to 0.45 meeting of the society a year, whereas there are many members who go to two or more.

Let us turn now from the aural and interpersonal communication at meetings to the recording of conferences in published proceedings. It is interesting to compare the rough estimates in Table XIV.17 for numbers of conferences of various types with the figures mentioned in Section 4.4 for published conference proceedings. The total of 1969 U.S. conferences and meetings listed in Table XIV.17 is 146; the total of conference proceedings listed in *Physics Abstracts* for 1969 is, after subtracting duplications, about 52 for conferences in the United States, 47 for conferences elsewhere. While these two sets of figures are not really comparable with each other, as they refer to rather different time periods and may well take in different amounts of material on the periphery of physics, they are consistent with the plausible speculation that regular national and regional meetings of societies do not usually publish proceedings (as distinguished from abstract bulletins) but that half or more of all other conferences do. A little less than half the proceedings, but rather less than a third of the total of papers, appear in books, the rest in journals.

Should conference proceedings be published at all, and if so, should they be published in books or in journals? The proper answer will not always be the same to either question. Publication does have a detrimental effect on the spontaneity of a meeting, and it clutters up the literature with many papers that are not in as suitable a form as they would be if published with fewer constraints on time and space and that sometimes greatly overlap other published work of the same authors. On the other hand, published proceedings provide a valuable record for participants, and in the best cases supply a handy means for the stranger to a field to get a bird's eye view of what is going on in it. If proceedings are published, two considerations are vital in determining the medium: Publication must be rapid, as the material is apt to be even more ephemeral than most; accessibility to nonparticipants must be maximized. The latter goal is, as we have indicated in Section 4.4, usually better achieved by publication in a journal.

6.2 COLLOQUIA AND SEMINARS

Like meetings, the category "lectures and classes" does not stand out very well above the fluctuations in the network of useful information channels in Figure XIV.7; its role in Figure XIV.4 seems fairly small, and in Figure XIV.5 it is unclear. The number in Figure XIV.3 is fairly certain, though measured only on physicists in a single institution: They spend an average of about 2.4 hours per week in seminars of various types in their own institution. This finding can be compared with an average of seven to eight hours per week of seminars announced in physics, plus others in chemistry, mathematics, engineering, and the like. It is clear that in large institutions the physicist's decision whether to go to a talk is based on a competition between available time and the interest of the topic announced. In small institutions, on the other hand, the number of talks offered may become small enough that most of the physicists go to all of them. In large universities, the time devoted to seminars may well be greater (though not by a large factor) because of sociological and educational pressures.

Suppose, then, that for want of a better estimate we accept something like 1.5 hours per week, or 75 hours a year, as the time the average physicist devotes to seminars. What can we conclude? One easy comparison to make is with the time spent listening to lectures at meetings. The latter time probably averages to no more than 10–15 hours per man-meeting per year. So it is clear that very few physicists spend as much time listening to speakers at meetings as at home; for the average physicist, the ratio must be of the order of one fifth or less. Now in most large universities, and probably in many other organizations, a majority of the speakers at seminars are visitors from outside the host institution. There seems to be no reason why, in these cases, the benefits to the listeners should be any less than for a talk at a meeting; in fact, they should be rather more, as there is normally much more time for discussion and for conferring afterward. The latter factors also enhance the benefits the speakers receive from audience response but may be offset by the fact that a talk at a meeting can draw experts from many institutions. In sum, it seems clear that the total utility of communications via seminar talks from outside speakers must be considerably greater than that via talks at meetings, since a much larger volume of listening is involved.

The custom of inviting outside speakers to seminars has an enormous side benefit. During the visit, all sorts of topics, often quite unrelated to the talk, come up for discussion between guest and hosts. Thus seminars contribute indirectly but very importantly to oral personal communication, which, as we have seen in Figures XIV.3 to XIV.7, is one of the two leading media for transmission of useful information.

This is wonderful; what does it all cost? Very little, except for the time of the participants, which we must let them evaluate for themselves. Examination of travel records suggests that typical institutions spend about half as much for seminar-speaker and miscellaneous consulting travel as for attendance at domestic meetings. With the estimate given in Section 6.1 for the total cost of the latter, we get an annual investment of the order of \$3 million to \$3.5 million for seminar talks and mutual consultation by U.S. physicists. This estimate seems reasonable in terms of another approach: One could get the same number by assuming that about half the physicists in the country each spend $2/3 \times 1.5 = 1$ hour a week listening in groups of 20 to 30 to outside speakers, and that these speakers typically come from nearby locations for which travel costs are only half as great as for the average trip to a meeting.

The reason, of course, that one gets more benefit for less money than in the case of talks at meetings is that in the one case, speakers and listeners both travel, while in the other case, only the speakers do so. But this does not mean that individual visits can serve as a substitute for meetings. In both cases, we have seen that a very large part of the benefit comes from the informal personal contacts that are generated, rather than from the talks. But for the seminar speaker, these contacts are all between the speaker and the staff of the host institution, whereas at a meeting many contacts are generated between people who are not speakers or who would not be likely to be widely invited as seminar speakers. Also, contacts at meetings are often the forerunners of seminar speaking visits. As for the talks themselves, while they unquestionably communicate more information per dollar when given as seminar talks by visitors to institutions than when given at meetings, it is clearly not practical to have each of the myriad of institutions visited by any sizable fraction of the limited number of speakers people want to hear. In other words, the value of the speakers' time becomes a limitation when the benefit they themselves receive from the visits begins to reach the saturation point.

Two further features of seminars deserve comment and apply as much to internal seminars as to those with outside speakers. One is the general sharpening of wits that accompanies discussion in a fair-sized group and the incentive that such discussion, or the threat of it, provides for clarification of one's grasp on a problem. The other is the education role of seminars, a topic we eschew in deference to the Report of the Panel of Physics in Education, and their role in interdisciplinary communication, which we shall take up in Chapter 8.

6.3 PERSONAL ORAL COMMUNICATION

We have seen much evidence in Chapter 2, especially Figures XIV.3 to XIV.7, that personal oral communication is by far the most extensively used medium of communication for physicists and that it is second only to the primary literature in terms of the amount and value of the information it communicates. As the value of the time invested in this sort of communication is enormous, it behooves us to try to understand as much as we can about how it operates, how it couples to other communication media, and on what its effectiveness depends. Unfortunately, among the studies that have been made in this field, few have dealt explicitly with physics, so we shall be content here merely to quote a few highlights from studies in other fields and to make a few general remarks.

One of the most obvious questions concerns the relative roles of intra-organizational and interorganizational conversations. The former are, of course, far more numerous, but the latter, when they occur, are more likely to involve persons with vital information to communicate. The study of high-energy theorists by Libbey and Zaltman[18] to which we have referred in Sections 2.1 and 4.6, has some illustrative statistics, shown in Table XIV.18. The figures on numbers of episodes look a little lower than one would infer from the expectation (Figure XIV.3) that seven to eight hours per week should be involved, and the fraction of the episodes involving people from other institutions is surprisingly high. It is hard to see how

TABLE XIV.18 Use and Value of Various Types of Personal Oral Communication by about 600 High-Energy Theorists

Type of Communication	Average Number of Communication Episodes in Two-Week Period (Standard Deviation)	Figures of Merit[a]
Face-to-face discussions with colleagues working at own institution	8.9 (11.4)	.07, .21, .23
Face-to-face discussions with colleagues not attached to own institution		.08, .14, .20
Conducted at own institution	2.4 (5.0)	
Conducted elsewhere	1.7 (3.1)	
Telephone conversations		.004, .011, .17
Received	1.6 (5.2)	
Initiated with party outside own institution	1.3 (4.8)	

Source: Libbey and Zaltman.[18]

[a] Same three figures defined in Table XIV.5, note *c* (first two based on percentage of replies rating a given channel as the best for some purpose; the third figure, based on a composite of numerical value ratings).

there could be over-reporting of the number of the latter episodes, however. We conclude:

1. Conversations with colleagues from other institutions are really quite frequent, at least in some areas of physics.
2. For high-energy theory, at least, the figures of merit all seem to indicate that such conversations are considered to be of comparable total value to those with colleagues at one's own institution.

These results underline the remarks made in connection with invited seminar speakers in Section 6.2. It is interesting to note the wide disparity among the three figures of merit for telephone communication; apparently few people give it top rating, but most consider it of fair importance. We suspect that long-distance calls are less used than they deserve to be. There have been many studies of interpersonal and other communication in broader populations of scientists and engineers, especially in industrial organizations. We shall not try to discuss these in any detail, mentioning only a few tidbits from them. A study by Rosenbloom and Wolek[104] of engineers and scientists in a number of industrial organizations indicates that point 2 is valid for the scientists in these organizations (16 percent of data sources oral external, 27 percent oral internal) but not for the *engi-*

neers (11 percent to 51 percent). The Columbia University study of polymer chemists[22] also supported 2.

3. The relative use of conversations with external colleagues increases with increasing educational level and decreases as one passes from research to development to operation.[104,105]

4. The use of conversations with external colleagues is positively correlated with attendance at meetings.[104]

5. Much of the information that scientists and engineers receive through personal contacts is information that the receiver would not have thought of seeking but that nevertheless turns out to be very useful.[22,104]

The low average use of extraorganizational sources by engineers and applied scientists may not mean, however, that such sources are unimportant in the total chain of communication of information to them, just as their low average use of literature does not mean that literature is unimportant. As Allen and his collaborators have shown,[26,36,37] information from outside an industrial organization typically penetrates it through special individuals, called "technological gatekeepers." These few individuals, far more than the rest of the staff, maintain personal contacts with outside persons and read external literature; they are then used as internal sources by their colleagues. (It would be interesting if studies could be made to see how important this concept is in a typical university environment.)

In general, as we have seen in Figure XIV.7, information flows in a network of channels, connected now in series, now in parallel. Large use of any one type of channel by many of the individuals of a group certainly means that this type of channel is important to the group; but low *average* use may mean only that use of this type of channel is centralized in gatekeepers, for whom (hence for the group as a whole) it may still be very important. The high use of personal oral communication means that it is surely very important; but without nourishment by journals, books, preprints, talks, and the like, it would dry up.

So much for the roles of different kinds of oral personal communication. What can be done to improve their effectiveness? Possible measures or policies are of two types:

1. Organizational. Oral communication can be affected by such things as patterns of responsibilities of members of an organization, transfers of personnel between different organizations and suborganizations, and ease of arranging time and funds for travel and for invited speakers. In regard to travel to meetings, there should be wider recognition of the fact that the value received by an organization from attendance at a meeting by one of

its staff members is dependent only to a minor degree on whether or not he gives a paper.

2. Physical and geographical. The strong influence of physical proximity should be borne in mind. It seems to be true that interdisciplinary communication (for example, physicists with chemists) is usually better in industrial laboratories than in universities, and this may be largely due to their all being under the same roof. Meetings should be planned with ample space and time for informal contacts.

Interdisciplinary communication, to which we have just referred, is apt to be especially dependent not only on talks, as noted earlier, but also on personal contacts. We shall take up this dependence in Chapter 8.

6.4 PERSONAL LETTERS

Personal written correspondence has been even less studied than any of the other types of informal communication. According to Figures XIV.3, XIV.4, and XIV.7 of Chapter 2, its role is rather smaller than that of oral conversations, but its yield of useful information is high in proportion to the time spent on it. Letters are obviously especially important for communication with colleagues abroad. As the writing of letters is a very individual matter, it is difficult to make any recommendations for improving this communication channel, other than to note that its effectiveness doubtless depends very much on personal acquaintanceships arising from international meetings and international visits; this factor should be borne in mind in assessing the value of such meetings. It would be helpful if a few studies could be done on the types of communication by letter and on some of the difficulties sometimes encountered.

7 Information Analysis Centers

Chapter 5 was devoted to the consolidation of information and discussed several kinds of consolidation. But one important kind was omitted, which we shall now consider: the information analysis center. This is an organization, large or small, whose primary purpose is to sift useful information (usually though not necessarily numerical data) from the primary literature, evaluate it, order it, and repackage it. Clearly, if enough people, and enough at the highest level of scientific competence, were employed at such centers, the problem of consolidating information could be solved in a very orderly and satisfactory way. The Weinberg report[1] in fact envisioned such well-staffed centers as the hope of the future, with the words, "Ultimately we believe the specialized center will become the accepted retailer of information, switching, interpreting, and otherwise processing information from the large wholesale depositories and archival journals to the individual user."

How far have we come toward this goal, and what are the prospects for the future? Let us first look at a few facts; perhaps then we can make a modified prospectus. According to a recent government listing,[106] there were in 1968 some 113 federally financed information analysis centers in the natural and social sciences. While it is sometimes estimated that the total number of such centers, with or without federal support, is several times larger than this, it is noteworthy that an investigation of some 400 centers by a COSATI panel[107] revealed that only 111 of them seemed to fall under their definition of information analysis centers. In physics, particularly, it is unlikely that there are many such centers without federal support. As Figure XIV.48 shows, the number of federally supported centers in all fields has been growing rapidly; this growth no doubt influenced, and has in turn been stimulated by, the commendation of such centers in the Weinberg report.

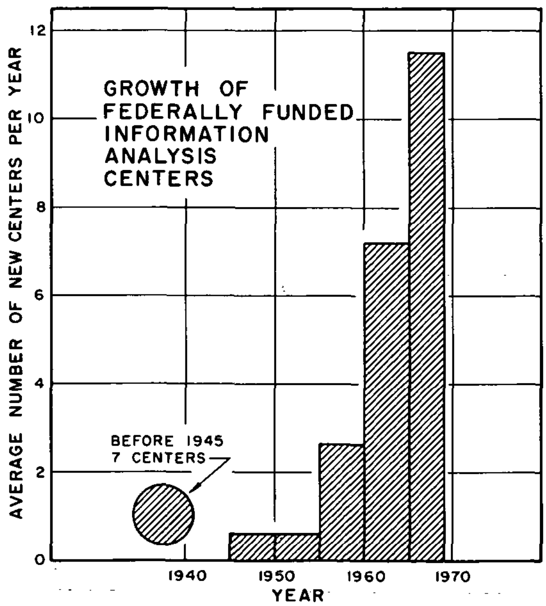


FIGURE XIV.48 Growth of federally funded information analysis centers (in all fields). Heights of bars are mean numbers of centers existing in 1968 that were started per year over the interval spanned by the widths of the bars; areas of bars are thus contributions to the total number of centers.[106]

These statistics are for centers in all fields, many of them in the domain of social sciences. Inspection of the list of centers[106] reveals, however, that quite a large proportion of them are in areas normally considered to belong to physics, and that many more are in areas of chemistry, geophysics, meteorology, space science, engineering, and the like that border on physics. To give a more definite flavor to the statistics, we have listed in Table XIV.19 some 29 federally financed centers whose work is perhaps closer to physics than to any other of the traditional disciplines. It is characteristic of the growth of the field that a more recent listing[108] gives some 34 in the same areas. Several broad features are evident in the table:

1. Eleven of the centers are operated by the National Bureau of Standards and six more are located elsewhere with NBS support. These (59 percent of the table) are components of the National Standard Reference Data System (see discussion below).
2. The largest number of centers—11 to 13—pertain primarily to solid-state physics; some 6 to 8 pertain to atomic and molecular physics; about 6 pertain to nuclear physics. One is in particle physics; one is primarily devoted to instrumentation.

Examination of the detailed descriptions of these physics centers[106] reveals further:

3. Nearly all of them deal primarily with numerical data, though many of them also prepare bibliographies and critical reviews. Many will undertake literature searches and other special tasks on request from outside scientists or organizations; they may charge the requestors for those services.
4. Most of the centers are small groups of one to a few physicists with a little clerical, bibliographic, and other support. However, they are usually located in large institutions where there is a considerable reservoir of expertise in the field of the center. Two or three centers seem to have over ten equivalent full-time scientists.
5. About three quarters of the centers seem to be primarily discipline-oriented, the remainder primarily mission-oriented.

Since we have seen in point 1 that a majority of the centers, including many of the most important ones, are run or supported by the Office of Standard Reference Data, it is appropriate to say a few words about this organization. (For additional details, see National Bureau of Standards.[108]) This agency was set up in the National Bureau of Standards in 1963 on the initiative of the Federal Council for Science and Technology; its powers

TABLE XIV.19 Selected Federally Financed Information Analysis Centers Most Closely Related to Physics[a]

Name of Center	Location[b]	Sponsor[b]	Scope (physics part)
Alloy Data Center	NBS	NBS	Structure-insensitive properties of metals and alloys
Applied Science Data Group	LBL	LBL	Neutron and photon cross-section data
Atomic and Molecular Processes Information Center	ORNL	NBS, AEC	Atomic collisions, atom-surface collisions
Atomic Energy Levels Data and Information Center	NBS	NBS	Atomic spectra
Atomic Transition Probabilities Data Center	NBS	NBS, ARPA	Atomic transition probabilities
Berkeley Particle Data Center	LBL	LBL, CERN	Properties of elementary particles and resonant states
Center for Diffusion in Gases	U. Md.	NASA	Gaseous transport data
Charged-Particle Cross-Section Information Center	ORNL	AEC	Charged-particle-induced reaction cross sections
Cryogenic Data Center	NBS, Boulder	NBS, NASA	Low-temperature properties of materials
Crystal Data Center	NBS	NBS, ACA	Crystal structure data
Diatomic Molecule Spectra and Energy Levels	NBS	NBS	Spectra of diatomic molecules
Diffusion in Metals and Alloys Data Center	NBS	NBS	Diffusion in metals and alloys
Electronic Properties Information Center	Hughes	USAF	Electronic, magnetic, and optical properties of solids
High Pressure Data Center	BYU	NBS	High-pressure properties of materials
Infrared Information and Analysis Center	U. Mich.	ONR	Infrared technology
Isotopes Information Center	ORNL	AEC	Production and use of radioisotopes
Soviet Institute for Laboratory Astrophysics Information Center	U. Colo.	NBS, ARPA	Collisions between electrons, photons, ions, atoms, and molecules

Light Scattering Data Center	Clarkson Coll.	NBS	Light scattering in gases and liquids
Low Temperature Specific Heats	NBS	NBS	Heat capacities 0–300°K
Microwave Spectra	NBS	NBS	Microwave absorption spectra of molecules
Nuclear Data Project	ORNL	AEC	Nuclear levels
Photonuclear Data Center	NBS	NBS	Photonuclear cross sections etc.
Radiation Effects Information Center	Battelle	NASA, DASA	Effects of radiation on materials
Radiation Shielding Information Center	ORNL	AEC, DASA	Shielding of reactors, accelerators, radioisotopes
Research Materials Information Center	ORNL	AEC	Availability, preparation and electronic properties of high-purity inorganic solids
Shock Wave Data Center	LBL	LBL	Hugonoit curves for chemically classifiable materials
Superconductive Materials Data Center	GE	NBS	Properties of superconducting materials
Thermophysical Properties Research Center	Purdue	NBS, USAF, NASA, NSF, ONR, and industry	Thermodynamic and thermal transport properties of materials and surfaces
X-Ray Attenuation Coefficient Information Center	NBS	NBS	X-Ray attenuation data useful in shielding

[a] Data from COSATI.[106]

[b] Abbreviations: ACA, American Crystallographic Association; AEC, U.S. Atomic Energy Commission; ARPA, Advanced Research Projects Agency; Battelle, Battelle Memorial Institute; BYU, Brigham Young University; CERN, European Organization for Nuclear Research; DASA, Defense Atomic Support Agency; GE, General Electric Co.; Hughes, Hughes Aircraft Co.; LBL, Lawrence Berkeley Laboratory; NASA, National Aeronautics and Space Administration; NBS, National Bureau of Standards; ORNL, Oak Ridge National Laboratory; U. Colo., University of Colorado; U. Md., University of Maryland; U. Mich., University of Michigan; USAF, United States Air Force.

and responsibilities were later spelled out by Congress in the National Standard Reference Data Act of 1968. Since that time, it has had a direct congressional appropriation; its current annual budget is about $2.4 million. Besides supporting many in-house and external information analysis centers (about half in physics), it also stimulates and funds a limited number of short-term data-compilation and bibliographic projects. It receives many advisory services from the scientific community, which are focused through the Numerical Data Advisory Board of the National Academy of Sciences and National Research Council. It maintains liaison with similar groups in other leading countries and with CODATA, the Committee on Data for Science and Technology of the International Council of Scientific Unions.

While we have seen in 3 above that most of the work of the present centers is on evaluation and compilation of numerical data, centers of other types are perfectly possible, and, indeed, were clearly envisioned in the quotation from the Weinberg report in the opening paragraph of this section. The compaction of information on concepts and relationships is apt to require, however, a rather different kind of organization from that used for the gathering and evaluation of taxonomic data. In particular, to adequately recognize and evaluate the many interrelationships among the many narrow specialties, a rather large staff would ordinarily be required. To review adequately a field of any breadth, many papers would have to be evaluated and interrelated, and the necessary expertise could often not be found in any small number of people or assimilated from offhand conversations with staff not attached to the center.

Still, it is not impossible to set up large continuing centers devoted to review work, although it has not been done in physics. The experience of the Gmelin-Institut, in inorganic chemistry, provides an example of the possibilities and the difficulties of such large ventures.[109,110] This institute, which has been operating in one form or another for a century and a half, undertakes to prepare continually updated versions of *Gmelin Handbuch der anorganischen Chemie,* based on a critical synthesis of the entire existing literature pertaining to all inorganic elements and compounds. The volume of literature reviewed and the amount of information extracted from it are impressive—even as far into physics as the semiconducting properties of germanium. (Many physicists are probably unaware of the possible utility of this *Handbuch* to them.) The intellectual work is done by some 80 scientists, with a comparable supporting staff. But the compaction process is slow, and it may be decades before a given volume is replaced by a new edition.

Far less ambitious, but still going beyond the present data-compilation centers, is Goudsmit's[90] (private communication, 1971), idea of a review-writing center. Such a center, which could be fairly small in terms of permanent staff, would be located at a large institution doing leading work in a certain field. Its primary function would be to assist senior staff or (usually) visiting scientists in the preparation of reviews in this field by providing them with intellectual backup, including professional-level assistance, bibliographic and clerical services, and assistance in writing and editorial work. This assistance would encourage competent authorities to undertake such work in a summer, semester, or year of leave from their regular jobs and would improve both the scholarliness and the readability of the product.

What are the future prospects for information analysis centers in physics? Undoubtedly, their number, scope, and use should and will continue to expand. In a field like physics, however, in which most of the work is not taxonomic in its orientation, we feel that centers of the types now in existence will never be able to satisfy more than a fraction—albeit an important fraction—of the need for consolidation of information. And it is too distant an extrapolation with the future to predict whether centers of the Gmelin or the Goudsmit type can ever do so. Certainly, the Goudsmit type of center should be tried, as it has much to recommend it and can be tested without a vast investment of resources.

For both existing and new types of centers, the real question that must be answered is how valuable they are in comparison with other activities that might compete with them for funds. In some fields (for example, nuclear data[92]), the evidence is overwhelming that intensified work of the type now done by some of the centers is badly needed (though some of it can perhaps be done more effectively by the short-term decentralized projects described in Section 5.7). In other fields operational-research studies may be needed. In dealing with questions of value and alternative modes of support, one must bear in mind the importance of ready accessibility of the products of a center to potential users. Thus:

1. The users need to be made fully aware that the products or services can be obtained and of how to find them. Channeling of this output into a few standard publications—such as *Nuclear Data* and the new AIP-ACS-NBS *Journal of Physical and Chemical Reference Data*—could be a great help.

2. Where the same product—a table of data, for instance—can be of use to many different people or organizations, it should ordinarily be offered to them at a price close to the cost of replicating the product once more rather than a price designed to recoup the entire production cost. This statement is a consequence of the general economic principle enumerated in Section 4.2 in connection with journal economics. This principle seems to have been pointed out in the congressional testimony associated with the National Standard Reference Data Act.[108]

With so many centers in existence and more springing up, one is entitled to worry about coordination. Fortunately, the fact that a majority of the existing centers are supported by the Office of Standard Reference Data makes possible a coordination of their work through this Office. In other cases, some cooperation among different agencies, and perhaps an expanded role for scientific societies, may be called for.

If information analysis centers are to play a growing role in the compaction of information, it will be necessary for them to recruit at least part of their staffs from among the most talented scientists. These will rarely be willing to work full time at the compaction of information; indeed, if they are to do this job well, they must keep in contact with the research front by remaining active in research. Thus a key recommendation must be that *information analysis centers, especially if they deal to any extent with conceptual information, should allow their leading staff members to apportion their time reasonably between new research and consolidation work.*

8 Communication of Physics with Other Disciplines and with the Public

So far, we have given most of our attention to channels used by physicists in communicating with each other. These channels are indeed very important: Without them, we could not have criticism of new ideas and development of the consensus that makes physics a science. But the notion that the nation as a whole should support physics must rest on its communication with the rest of our culture and with technology. Therefore, we must understand how much communication of this sort there is, through what channels it takes place, and what determines its effectiveness.

Despite the importance of the subject, this section will be short, because the main communication channel for it, and certainly the one about which the most statistics are available, falls outside our purview. We refer, of course, to formal education, whose role in communication of physics information to scientists in other fields, to engineers and technicians, and to the public is discussed at length in Chapter 11 of Volume 1 of *Physics in Perspective* and the Report of the Panel on Physics in Education. Interdisciplinary education is also touched on in the reports of several of the other Panels. Although we shall mention formal education in several places, we shall not discuss it in detail. Our program will be to start with a discussion of the remaining channels for dialogue between physics and its sister sciences and engineering and continue with a few words about communication with the other components of our culture.

8.1 COMMUNICATION WITH OTHER SCIENTIFIC AND TECHNOLOGICAL DISCIPLINES

8.1.1 *Primary Literature*

Research papers in any scientific field are usually not readily understood by people trained only in a different discipline. The language of such papers generally is comprehensible only to people trained in the field or subfield of the research. Nevertheless, workers in specialties not too far beyond the boundaries of physics can often make direct use of papers somewhat on the physics side of the boundary and vice versa. Citation matrices of the sort shown in Figure XIV.29 can be very illuminating, although this particular figure has few data on journals outside of physics. Some counts of citations in journals of leading societies in nonphysics fields have been made (see, for example, Herschmann[111]). These have shown that references to publications of the American Institute of Physics constitute 3.7 percent of the citations in journals of the American Chemical Society, 8.2 percent of those in journals of the Institution of Electrical and Electronics Engineers, and 4 percent of those in journals of the American Nuclear Society. References to all physics journals in the world might well be twice these numbers. Thus, even the primary physics literature has a significant direct effect on neighboring fields.

An alternative way of assessing interdisciplinary communication in the primary literature is to look for journals likely to have readers or contributors in two or more different disciplines. Among such journals would be nearly all those between the two curves in Figure XIV.17, plus some of the "physics" journals. As an example of this approach, we have examined the 78 leading journals that Keenan and Brickwedde[49] ranked in order of the number of items they contributed to *Physics Abstracts* in 1965. (These were all that contributed more than 0.25 percent of the entries; their total contribution was 82 percent of the entries.) Table XIV.20 lists those journals that might be considered interdisciplinary, roughly classified into five types according to whether they were truely multidisciplinary, were on the near side or the far side of the boundary between physics and some other major discipline, intentionally bridged such a boundary, or were devoted to a fairly narrow but interdisciplinary specialty. It will be seen that the journals listed in the table account for nearly 29 percent of *Physics Abstracts* entries, a respectable portion of the 82 percent covered by all journals to this level.

TABLE XIV.20 Journals with a Significant Interdisciplinary Role That Contributed More than 0.25 Percent of the Entries in *Physics Abstracts* 1965

Multidisciplinary[a]	Centered in Physics but Near a Boundary	Bridging Physics and a Neighboring Discipline	Centered in Neighboring Discipline but Near Physics	Devoted to a Narrow but Interdisciplinary Specialty
Nature	J. Appl. Phys.	J. Chem. Phys.	J. Geophys. Res.	J. Am. Ceram. Soc.
J. Res., Nat. Bur. Stand.	Rev. Sci. Instrum.	Fiz. Met. Metalloved.	Trans. AIME	Atomnaya Energiya
Proc. Roy. Soc. (London) A	Prib. Tekh. Eksp.	Nucl. Sci. Eng.	Astrophys. J.	Planet. Space Sci.
Science	Appl. Opt.	Acta Met.	J. Atmos. Terrest. Phys.	Surface Sci.
	Zh. Tekh. Fiz.	Solid State Electron.	Radiotekh. Elektron.	
	J. Acoust. Soc. Am.		J. Phys. Chem.	J. Mol. Spectron.
	J. Opt. Soc. Am.		Proc. IEEE	Teplofiz. Vys. Temp.
	J. Sci. Instrum.			
	Nucl. Instrum. Methods			
	Brit. J. Appl. Phys.			
	Z. Angew. Phys.			
	Jap. J. Appl. Phys.			
	Appl. Phys. Lett.			
	Ind. J. Pure Appl. Phys.			
	Akust. Zh.			
2.7%[b]	16.0%[b]	7.5%[b]	7.0%[b]	2.4%[b]

[a] Interpreted as containing papers from two or more major disciplines without complete segregation into different sections of the journal.
[b] Percent of all entries from all journals in column.

Particularly important are the journals of the middle column, which consciously try to span a wide range of topics from one side of the interdisciplinary barrier to the other. The range is particularly great for a journal like *Fizika Metallov i Metallovedenie,* which publishes papers on topics ranging from abstruse quantum-mechanical, many-body theory to metallographic studies of aging of steels. Presumably, such journals operate in the hope that scientists well on one side of the interdisciplinary boundary will, in perusing the journal for material in their specialty, occasionally be intrigued by material from the other side of the line.

8.1.2 *Books and Reviews*

What about books and reviews? Logically, these would seem to be the best means for transmitting written information among disciplines, as in them the information from one discipline can be gathered together, sifted for its relevance to a second discipline, and presented in language that will be comprehensible to workers in this discipline. Much such writing does indeed take place. In many cases, the short review article is the most appropriate medium, as the worker in the second discipline often cannot afford to spend much time learning about a topic in the first. But even books directed across disciplinary lines are numerous and extensively used. Some illustrative figures are given in Table XIV.21: The entries show the numbers of books, in a random sample from the physics shelves of a research library devoted to physical science and technology, that were addressed primarily to various classes of readers, as determined from statements in the prefaces. Only 34 percent of the sample was aimed primarily at research workers in physics, and another 11 percent to 17 percent at educating those who have committed themselves to physics; 39 percent to 45 percent were primarily for education of or use by workers in other fields; and 10 percent were for education of or use by people who do not necessarily have a career interest in science or engineering.

Thus, there is a fairly large amount of literature that undertakes to repackage physics information for workers in other fields. It would be useful to know how effective and how valuable it is, but we are not aware of any studies on these questions.

8.1.3 *Conferences*

One would expect oral communication to be even more useful for the crossing of disciplinary lines than within a discipline, at least if the right people can be brought together. It takes less courage and energy to listen to someone or to converse with him than to tackle an unfamiliar book or

TABLE XIV.21 Intended Primary Readership of a Random Sample of Books (Published 1960-1970) in the Physics Section (Dewey Decimal Numbers in the 530's) of a Research Library Devoted to Physical Science and Technology

Principal Use Envisioned	No. of Books	Typical Titles
High school or technical school[a]	4	
Scientifically literate public[a]	10	Nuclear Energy; The Structure of Atoms
Undergraduates, other than physics majors[b]	12	Modern Electromagnetic Fields
Undergraduate physics majors	7	
Graduate students in physics	13	Atomic Spectra; Symmetry Properties and Fundamental Particles
Graduate students in nonphysics fields[b]	4	Introduction to Quantum Electronics
Both types of graduate students[b]	10	
Research physicists[b]	60	Plasma Dynamics; Cosmic Radiation and High Energy Interactions
Research or practicing engineers or applied physicists[b]	25	Fundamentals of Infrared Technology; Dynamic Behavior of Thermoelectric Devices
Research chemists or chemical physicists[b]	13	Silicon Carbide; Progress in NMR Spectroscopy
Research workers in or bordering on biology, astronomy, mathematics, or space science[b]	16	Electromagnetodynamics of Fluids; Speech Analysis, Synthesis, and Perception
Miscellaneous philosophical, historical, or cultural[a]	4	
	178	

[a] Communicate with nonscientists.
[b] Interdisciplinary use likely.

article, and mismatches of vocabulary and ways of thinking can be overcome by questions and answers. Interdisciplinary conferences provide just these opportunities. We saw in Table XIV.19 that there are more special-topic conferences than any other sort, most of them one-of-a-kind conferences, but a few periodic. Many of these conferences are on topics that cut across disciplinary lines, and these provide splendid opportunities for the exchange of ideas between physicists and other groups such as chemists, engineers, metallurgists, mathematicians, and biologists.

Best of all are those of the Gordon Research Conferences that are in a position to maintain a balanced interdisciplinary participation. These conferences have several great advantages: They have review talks that are sufficiently unhurried to permit clear development of a subject and ample questioning; they allow a great deal of time for informal personal contacts;

they are repeated each year on closely related topics, so that acquaintances formed in one year can be renewed in a subsequent one. However, these benefits are achieved only by making the conferences small; thus, only a limited number of especially active workers in the two or more disciplines involved can reap the benefits.

8.1.4 *Seminars*

Very similar in effect to conferences attended by people from many organizations are seminars for the staff of a single institution. If the pattern of seminars is rigidly partitioned along departmental lines, the physicists may rarely attend the chemists' or the geologists' seminars, and vice versa; but with the frequent choice of topics of interdisciplinary interest, plus a little care in scheduling and announcements of joint sponsorship, significant communication among disciplines can often be achieved.

8.1.5 *Interpersonal Communication within Organizations*

Such communication is another very effective type of oral interdisciplinary communication. It is fostered in industrial laboratories, where workers trained in different disciplines are often close together physically or even organizationally (as they usually are not in universities). It is not uncommon for a physics PhD to start working in the basic-research area of an industrial laboratory and later to transfer to a development area, which is stimulated by his new perspective and approach. Less common, but sometimes important, are transfers into physics areas by chemists, mathematicians, and representatives of other disciplines. In universities, interdisciplinary departments and institutes can often be effective, as can interdisciplinary degree programs and thesis projects. These have disadvantages as well as advantages; however, their discussion belongs in the realm of education, which is outside our province here.

In closing this subsection, we should like to call attention to a revealing technique for measuring the intellectual penetration of one field into another that has not yet been applied to physics but whose application to the chemistry–mathematics relationship provides a nice example of the way all the sciences are becoming more interrelated. Orient[112] has counted the frequency of occurrence, in Soviet articles on analytical chemistry, of words and phrases from the jargon of mathematical statistics. This frequency has risen spectacularly during the period 1961–1967, and this rise can plausibly be correlated with the influence of certain interdisciplinary books and articles.

8.2 COMMUNICATION WITH LEADERS IN CULTURE AND PUBLIC AFFAIRS AND WITH THE GENERAL PUBLIC

Science is too important to be left to scientists. Its technological progeny must be managed in the public interest; its intellectual implications extend far beyond specialist boundaries. From Lippert[113] we borrow a quotation from Einstein:

Die Beschränkung der wissenschaftlichen Erkenntnisse auf eine kleine Gruppe von Menschen, auf die Spezialisten, schwächt den philosophischen Geist eines Volkes und führt zu dessen geistiger Verarmung.*

For both practical and cultural reasons, then, it is important that the public in general, and most especially its intellectual leaders and men of affairs, not be isolated from the progress of physics. Communication with these groups, even more than with any other, must rely on the public schools and the institutions of higher education. Outside of these, the existing communication channels are rather meager. There are the popular press and the intellectual magazines, books, and lectures. Personal contacts can be very effective, but there are not enough of them. On the other channels, we have only a few scattered comments.

8.2.1 *The Popular Press*

Reporting on scientific topics can be effective only if the reporters have some knowledge of and interest in science. Fortunately, the number and esprit de corps of journalists specializing in science writing seem to be growing slowly. There is a national organization, the Council for the Advancement of Science Writing, that assists working journalists, offers training programs, and conducts seminars for science writers. Some schools of journalism offer training in the field. The American Institute of Physics has a Public Relations Division that acts as middleman between reporters and physicists who present interesting new work in talks or articles.

Despite these favorable trends, the reporting of physics developments in even the leading newspapers and magazines is very uneven. Figure XIV.49 shows, for example, the coverage of a number of aspects of physics in *The New York Times* and in the *Reader's Guide to Periodical Literature* (which indexes articles from some 150–160 periodicals of wide circulation) for two periods about a decade apart. While, on a bulk basis, the magazines

*"The restriction of scientific insights to a small group of people, i.e., specialists, weakens the philosophical spirit of a people and leads to their spiritual impoverishment."

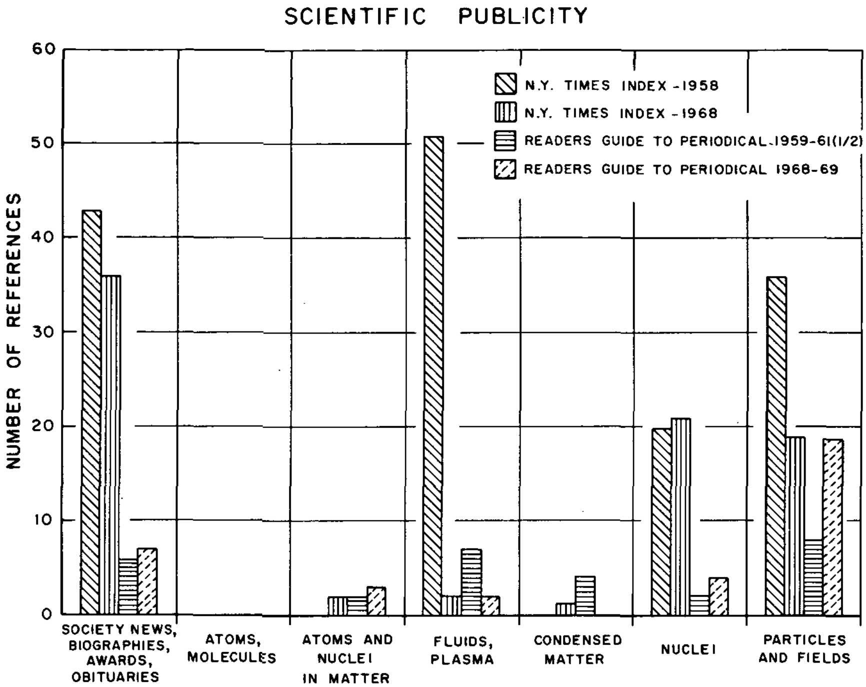


FIGURE XIV.49 Science publicity in *The New York Times* and leading periodicals.

have improved somewhat, the *Times* has not. One wonders why, since, during this decade, higher education has been expanding and the proportion of readers with college or graduate training is greater.

8.2.2 *Intellectual Magazines*

These are particularly important, since through them the world of physics can make contact with people who, on the one hand, have the intellectual capacity to apprehend profound ideas and the implications of new facts and who, on the other hand, are most likely to be in positions of leadership where any knowledge can be put to good use. Although much material in such magazines as *Science, Scientific American,* and the like can be, and sometimes is, read with profit by nonscientists, the great majority of the country's leaders in intellectual and practical affairs doubtless do not read them. Articles written with great skill in magazines of wider, but still intellectual, appeal could play a very useful role if the right physicists and scientific generalists would undertake them.

8.2.3 *Public Lectures and Television*

Lectures in a hall do not seem to be popular in the United States. By contrast, we have been told that in Kharkov, U.S.S.R., a public lecture series on the special theory of relativity about a decade ago drew an audience of about 500. Television is another matter. There have been some very successful programs for children and, in those localities where educational stations exist, there should be real possibilities for communicating to adults exciting and thought-provoking material about physics, as well as education of a more routine sort.

9 Roles and Responsibilities

In this final chapter, we shall focus, one at a time, on the various major groups that play roles relating to communication in physics in the United States, and, for each group, we shall review what has been said or implied in the previous chapters about its role and its responsibilities for future progress. Most of the groups we shall consider can be identified in Figure XIV.5.

1. Government agencies. These directly provide information services central to physics, at a level corresponding to perhaps a few million dollars a year and, in many cases, have large mission-oriented information programs peripheral to physics. More important for our present interests, however, they provide funds for many of the communication-related activities carried out by other groups. The dollar values of such activities are listed below.
2. Scientific societies. These are concerned with primary publication, secondary services, meetings, and the like, the total of budgets for these purposes (in physics) being of the order of \$8 million to \$9 million per year.
3. Commercial publishers and commercial information services. Remembering that most of the commercial journals in physics are foreign, as are some of the books, we can estimate that the portion of the work in Figure XIV.5 done by U.S. commercial organizations amounts to about \$7 million to \$8 million per year.
4. Research organizations. With their management of travel and their provision of library services, these oversee the lion's share of the resource

commitment that goes into physics communication, probably over $30 million per year. The other large component in the breakdown of activities in Figure XIV.5 is, of course, foreign organizations, whose work affecting the entries in the figure is worth approximately $10 million per year. We shall not discuss this component further here. There is one further group, however, that we should include as having roles and responsibilities, though it does not appear in Figure XIV.5, namely:

5. Individual physicists. Though individuals do not control large-scale activities, they make decisions that in the aggregate have a great influence on the effectiveness of the communication media operated by the larger groups.

9.1 GOVERNMENT AGENCIES

In several places in the previous chapters, we have encountered an important principle of economics that pertains to most kinds of information services and that practically requires a partial government subsidy of them if the public interest is truly to be served. The reader will recall that in discussing primary journals (Section 4.2) we argued that society is poorly served by pricing journals so as to recoup both the prerun and runoff costs. There will then be cases in which provision of a copy of the journal to an additional user will benefit his work, and through it society, by an amount greater than the cost of providing one more copy for him (the runoff cost) but less than the sum of runoff and prorated prerun costs. The positive net benefit cannot be realized if the journal is priced at the latter figure. The same general principle (which is well recognized by economists) applies to many other information services, for example, abstract journals, data compilations, and treatises. In each case, there is a product that is costly to produce initially but inexpensive to replicate. In each case, different prospective purchasers place widely different values on the product, but in each case, the purchaser's valuation is a plausible estimate of its value to society. Although a fully adequate treatment of what society gets for its money has to take account of such things as producer motivation and interactions between different users—these are discussed for the primary-journal case in SATCOM[4]—the essence of the argument as we have just sketched it remains valid.

As we have seen in Section 4.2, this incremental-pricing rule provides the major argument for the page-charge system of financing, as long as the funds for payment of page charges can be obtained from a source that has a long-range concern for the advancement of science. The policy of the Federal Council for Science and Technology[70] of encouraging the inclusion of page charges in research budgets reflects an intelligent concern in this direction, described in the statements that publication is an integral part of doing research and that those who support research have an obligation to support its publication. We commend continuation of this policy and urge further measures to ensure that the governmental contribution to page-charge support operates in the public interest. (See, for example, the discussion in SATCOM.[4])

Much the same is true for the basic abstracting and indexing services, whose role is just like that of primary publication, namely, to make possible effective public access to the results of research. While for reasons we have mentioned in Section 3.6, no techniques for subsidy of extragovernmental abstracting and indexing services have been worked out that are comparable in practicality to the page-charge system, we feel that agencies that support research should be aware that their work is incomplete as long as the basic secondary services that provide access to it have to recoup their entire costs from subscriptions. The mission agencies that support inhouse secondary services—most notably the Atomic Energy Commission with *Nuclear Science Abstracts*—are to be commended for their recognition of this principle. We agree, however, with the SATCOM Report[2] that basic secondary services for all physics should be managed by scientific societies rather than by mission-oriented federal agencies.

In the realm of compaction of information, the principle is again just as valid, and though far easier to apply, it seems to have been even less generally appreciated than for secondary services. We urge that all agencies supporting research accept a proportionate responsibility for consolidation of its results, that is, for funding the production of critical data compilations and reviews. As we have noted in Section 5.7 and Chapter 7, there are several measures that may be appropriate, most notably making available grants to individuals or small groups for review or compilation work and the funding of review-writing centers. When the need exists, support of ongoing information analysis centers is recommended. Charges for the products of these centers should be set at about distribution cost, at least in the case of products with a wide range of users.

The general principle of providing support adequate to allow marketing of information services at incremental or distribution cost does not especially imply the desirability of subsidy of preprint distribution. As we have seen in Section 5.6, a centralized preprint distribution of small volume might merit consideration in some cases but should be virtually self-supporting in its steady state.

Insofar as it is possible to plan an orderly response to the fluctuations in the funding of science that national economic conditions may require, it would be desirable to put a larger proportion of the budget into consolidation activities in so-called lean periods of very limited funding and a smaller proportion in periods of ample funding. It takes fewer dollars to support a given number of scientists in consolidation work than in research.

Before closing this section, we should mention the broad issue of copyright, particularly as it applies to replication and to computerized information systems. Though this issue is far broader than science and technology, its great importance for these areas, including, of course, physics, has been described at length in the SATCOM report[2]; the legislative and judicial branches of the government have a major responsibility to establish guidelines that will serve the public interest by recognizing the valid points that have been made on both sides of the issue. In the meantime, intelligent use of direct and indirect subsidies to encourage pricing at output cost can significantly ease dislocations resulting from copying. For example, page-charge support for commercial journals, though admittedly very difficult to implement practically, might be explored further.[4]

9.2 SCIENTIFIC SOCIETIES

As we have seen in Figure XIV.5, the largest role of the societies, in economic terms, has been the management and production of primary journals. In the case of the American Institute of Physics and its member societies, this work has been carried out quite effectively, and the low prices and high circulations of its journals are a commendable example to the other sciences. Improvements in the technology of journal production are continually being made. We might offer a few words of caution, however. The rapid growth in bulk of some of the journals implies the need for vigilance and imagination in making decisions about subdivision. The use of the two-track system to encourage honoring of page charges (Section 4.2), though generally more beneficial than harmful (see SATCOM,[4] Section IVB4, for a quantitative balance of pros and cons), is not in the best interest of society if the nonhonoring papers are subjected to long delays.

Prompt repackaged journals—the "user journals" described in Section 4.3[74]—could conceivably be not only the answer to the loss of utility that journals have suffered due to the decline in personal subscriptions but also an unprecedentedly effective current-awareness tool. We encourage experimentation with them.

Although swamped by costs resulting from primary-journal publication, the secondary services have in recent years occupied a major place in the thinking and planning of the American Institute of Physics. Their planned "National Information System for Physics and Astronomy"[114] envisions a highly integrated structure of primary literature, access services, and consolidation of information; the health of this structure, and of many other aspects of the nation's physics effort, is to be promoted by an ongoing program for information about the functioning of the physics and astronomy community. We have seen in Section 3.6 that economy and speed can be served by integration of several different secondary services with each other and with primary journal composition; these possibilities are being exploited, as are those of cooperative agreements with other organizations that generate inputs for secondary services.

The consolidation portion of this AIP program so far has two clearly enunciated components, both of which seem promising. One is a new journal for publication of data compilations, which should make many of them much more accessible than they have been in the past. The other is a proposal for a government-funded system of all-expense grants to authors of reviews or compilations[114] (see our Section 5.7). The AIP or its member societies could take the initiative in implementing some of the other measures mentioned in Section 5.7.

In conclusion, we urge most especially that the physics community support an adequate ongoing operational-research group at the American Institute of Physics to monitor data and perform studies relating to the functioning of the physics community, including problems of information and communication.

9.3 U.S. COMMERCIAL PUBLISHERS AND INFORMATION SERVICES

The vital role of commercial publishers as producers of physics books is not due merely to the fact that other groups produce so few of them. It is also due to their unique capability for recognizing the needs of special groups of users. As we have seen in Section 8.1, most of the advanced books produced on physics subjects are addressed to one kind or another of interdisciplinary readership. The interaction of commercially motivated managers with diverse would-be authors is a very effective mechanism for discovering the existence of an important class of users with special needs and getting the right sort of information package prepared for them.

In regard to the advanced review literature of pure physics, we would like to see a little more stress on the thorough and comprehensive types of review articles and treatises; however, we recognize that this emphasis will depend more on authors than on publishers. In addition, collections of reprints of important papers in a special field can be extremely useful. Intermediate between these two types of consolidation, it often can be helpful to reprint individual review articles from review journals when they are much in demand. We should like to encourage collaboration of journal publishers with publishers of books in ventures of these last two types.

We should like also to see greater emphasis on a small number of "Progress in . . ." series of review-article collections and less stress on the publication of conference proceedings in isolated books. As we explained in Section 6.1, we feel that these are usually more accessible if they are published in journals and at the same time made available bound separately.

Serious concern has been expressed over the squeeze on publishers of monographs due to rising costs and the decline of sales due to increasing specialization.[115] Fortunately, some publishers are reducing costs by turning to typewriter composition and offset printing; moreover, as we noted in Section 5.5, paperback editions, usually fairly satisfactory for the individual purchaser, are becoming increasingly available at prices of the order of half those of hard-cover editions. (In general, *for a given level of profit,* the welfare of society is enhanced by any scheme of differential pricing that can create a correlation between the prices paid by different users and the value of an information product to them.)

9.4 RESEARCH ORGANIZATIONS

Our remarks in Section 9.2 strongly endorsed the page-charge system of support for primary journals. To be successful, however, this system requires the cooperation of the organizations in which research takes place. We should like to stress that it is precisely these organizations' own interests that are at stake: The object of the page-charge system is to make journals easily accessible in as many libraries, sublibraries, group collections, and individual holdings as possible. Such accessibility is surely desired by these organizations as users of information and by their staff members in their roles as authors. We reiterate here the point briefly mentioned in Section 4.2, that a decision of a U.S. author to publish a paper in an expensive non-page-charge journal can sometimes actually cost more money, from the budgets of U.S. research institutions, than publication, with payment, in a page-charge journal. Use of typical numbers in Equation (4.1) shows, for example, that this can be the case when the price is above about $0.10 per kiloword, unless the U.S. circulation of the journal in question is below 750. This arithmetic, moreover, is based solely on the purchase cost of the journals and does not take account of the fact that library administration, binding, and storage costs, which typically are comparable with acquisition costs, are appreciably higher per kiloword for the smaller journals. Thus an organization that encourages its authors to publish in expensive journals is really asking its sister institutions to support its publication, and at a rate that may be more expensive than that in a page-charge journal.

In Section 9.1, we stressed that those who foster research have an obligation to accept a proportionate responsibility for its consolidation. This obligation applies as much to universities and to industrial basic-research groups as it does to government. Actually, the interests of both types of organizations would be advanced if they would more actively encourage their staffs to prepare books, reviews, and compilations. The educational value of such work is obvious. Since a primary motivation for performance of basic research in industrial laboratories is to place the laboratory in touch with research-front activity in the rest of the world, what could be better than an occasional comprehensive review? Faculty and administration must eliminate the notion that three research papers are a better ground for a promotion than two research papers and a review. Similarly, we urge research organizations of all types to be hospitable to information analysis and review-writing centers and to encourage their staffs to take leaves to work at such centers.

Research organizations have a special responsibility to foster informal communication and to facilitate the use of formal media by their staffs. We have commented briefly in Sections 4.2, 6.3, and 8.1 on organizational and administrative measures to foster oral communication, especially among workers in different disciplines. Both in this realm and in that of access to books, journals, and secondary services, we stress again the importance of proximity of the user to the resource (Section 2.1). Large buildings housing many workers in related fields and containing an adequate library in the middle are much to be preferred to a scattering of isolated buildings, some without libraries. Even within large buildings, subcollections of books and periodicals useful to a particular group can be most helpful. Research departments and libraries should cooperate closely and be aware of each other's problems.

9.5 INDIVIDUAL PHYSICISTS

Effective communication depends on the freedom of each individual to choose his degree of involvement with each of a wide range of information resources. It is only he who can know what he needs and how much. Unfortunately, several parts of our study (for example, Sections 3.5, 4.8, and 5.5) have shown that the physicist often does not know these things as well as he should and that, when he does know his needs, he often does not know how to go about satisfying them. The basic facts of the communication picture should be much more widely known. There should be less provincialism in journal-scanning habits. The sluggish response to new information tools—citation indexes, current-awareness services, translation journals—should be accelerated. Education in these areas should be a part of graduate training.

It is individuals, too, who undertake the preparation of books and reviews. We have made a general exhortation (Sections 5.6 and 5.7) for more and better work in this direction and have sketched specific desiderata for improving quality and readability. We reiterate, too, our recommendation, based on the citation statistics of Section 5.5, that authors not dissipate their review-writing energies in too many sketchy reviews and that they publish their reviews in media of high visibility and wide circulation.

Our remarks in Section 9.4 regarding the folly of publishing papers in expensive journals of low circulation should be taken to heart by authors as well as by administrators, as it is usually the authors who decide where to submit their work.

Last but far from least among the roles of individual physicists is the collective one of an informed and dedicated citizenry. It is their preferences that determine the demand for marketed products, their votes and letters that determine the policies of the societies; their aggregate pressure has a large though not necessarily controlling influence on the policies of governmental agencies and other organizations. We hope that, in examining these responsibilities, they will give serious attention to problems of information and communication.

Appendix A

Journals Contributing the Largest Number of Items to Each of the Subject Subdivisions of *Physics Abstracts* 1964,[4] Listed in Rank Order in Each Field. Also Given Are the Percentages of Total Entries in Each Section Supplied by the Journals Listed

A. *General*
(68 percent from 10 journals)

Am. J. Phys.
J. Sci. Instrum.
Rev. Sci. Instrum.
Prib. Tekh. Eksp.
Appl. Opt.
Nature
J. Phys.
Brit. J. Appl. Phys.
Zh. Vychisl. Math. Phys.
J. Res. Nat. Bur. Stand.

B. *Mathematical Physics*
(52 percent from 10 journals)

J. Math. Phys.
Phys. Rev.
Nuovo Cimento
C.R. Acad. Sci.
J. Chem. Phys.
Am. J. Phys.
Phys. Lett.
Ann. Phys. (N.Y.)
Zh. Eksp. Teor. Fiz.
Physica

C. *Mechanics*
(47 percent from 9 journals)

Am. J. Phys.
C.R. Acad. Sci.
Bull. Acad. Polon. Ser. Sci. Tech.
Prikl. Math. Mekh.
J. Franklin Inst.
J. Appl. Phys.
Abhandl. Dtsch. Akad. Wiss. Berlin
Dokl. Akad. Nauk SSSR
Z. Angew. Math. Phys.

D. *Fluids*
(37 percent from 10 journals)

J. Chem. Phys.
Phys. Fluids
J. Phys. Chem.
Rev. Sci. Instrum.
Ukr. Fiz. Zh.
C.R. Acad. Sci.
Z. Naturforsch.
Nature
Opt. Spektrosk.
Dokl. Akad. Nauk SSSR

E. *Vibrations, Waves, Acoustics*
(68 percent from 10 journals)

J. Acoust. Soc. Am.
Akust. Zh.
Acustica
Phys. Fluids
C.R. Acad. Sci.
Prikl. Mat. Mekh.
J. Sound Vibration
J. Appl. Phys.
Bull. Seismol. Soc. Am.
Zh. Vychisl. Math. Phys.

F. *Optics*
(61 percent from 10 journals)

Appl. Opt.
J. Opt. Soc. Am.
Opt. Spektrosk.
Rev. Sci. Instrum.
Optik
C.R. Acad. Sci.
J. Sci. Instrum.
Opt. Acta
J. Photogr. Sci.
Proc. IEEE

G. *Heat*
(46 percent from 10 journals)

Rev. Mod. Phys.
Phys. Rev.
Phys. Lett.
Zh. Eksp. Teor. Fiz.
Rev. Sci. Instrum.
Phys. Rev. Lett.
Int. J. Heat Mass Transfer
Inzh. Fiz. Zh.
J. Sci. Instrum.
J. Chem. Phys.

H. *Electricity and Magnetism*
(41 percent from 10 journals)

Zh. Tekh. Fiz.
Phys. Fluids
Prib. Tekh. Eksp.
J. Appl. Phys.
Rev. Sci. Instrum.
C.R. Acad. Sci.
Zh. Eksp. Teor. Fiz.
Phys. Rev.
Phys. Lett.
Appl. Phys. Lett.

I. *Nuclear Physics*
(66 percent from 10 journals)

Phys. Rev.
Nuclear Phys.
Phys. Lett.
Nuovo Cimento
Zh. Eksp. Teor. Fiz.
Nucl. Instrum. Methods
Phys. Rev. Lett.
At. Energ.
J. Phys.
Progr. Theor. Phys.

J. *Atomic and Molecular Physics*
(61 percent from 10 journals)

J. Chem. Phys.
Phys. Rev.
Opt. Spektrosk.
C.R. Acad. Sci.
Litov. Fiz. Sb.
J. Mol. Spectrosc.
Proc. Phys. Soc.
Z. Naturforsch.
J. Quant. Spectrosc. Radiat. Transfer
J. Math. Phys.

K. *Solid-State Physics*
(43 percent from 10 journals)

Phys. Rev.
J. Appl. Phys.
Fiz. Met. Metalloved.
Fiz. Tverd. Tela
J. Chem. Phys.
Phys. Status Solidi
C.R. Acad. Sci.
J. Phys. Soc. Japan
Acta Crystallogr.
Phys. Lett.

L. *Physical Chemistry*
(60 percent from 10 journals)

J. Chem. Phys.
Z. Naturforsch.
Rev. Sci. Instrum.
Nature
C.R. Acad. Sci.
Dokl. Akad. Nauk SSSR
J. Phys. Chem.
Fiz. Met. Metalloved.
Izv. Akad. Nauk SSSR Ser. Fiz.

M. *Geophysics*
(57 percent from 10 journals)

J. Geophys. Res.
J. Atmos. Terrest. Phys.
Planet. Space Sci.
Izv. Akad. Nauk SSSR Ser. Geofiz.
Nature
J. Atmos. Sci.
J. Res. Nat. Bur. Stand.
Can. J. Phys.
J. Acoust. Soc. Am.
Quart. J. Roy. Meteorol. Soc.

N. *Astrophysics*
(59 percent from 10 journals)

Astrophys. J.
Astron. Zh.
Nature
Monthly Not. Roy. Astron. Soc.
Astron. J.
C.R. Acad. Sci.
Ann. Astrophys.
Z. Astrophys.
J. Geophys. Res.
Planet. Space Sci.

O. *Biophysics, Physiological Physics*
(86 percent from 10 journals)

J. Acoust. Soc. Am.
J. Opt. Soc. Am.
Acustica
Vision Res.
Opt. Acta
Nature
Nucleonics
Rev. Opt.
Rev. Sci. Instrum.
Science

P. *Technique, Materials*
(84 percent from 7 journals)

Rev. Sci. Instrum.
J. Sci. Instrum.
Prib. Tekh. Eksp.
Brit. J. Appl. Phys.
Z. Angew. Phys.
Am. J. Phys.
J. Chem. Phys.

References

1. A. M. Weinberg, W. O. Baker, and K. Cohen, *et al.*, *Science, Government, and Information: A Report of the President's Science Advisory Committee* (U.S. Government Printing Office, Washington, D.C., 1963).
2. Committee on Scientific and Technical Communication, *Scientific and Technical Communication: A Pressing National Problem and Recommendations for Its Solution* (National Academy of Sciences, Washington, D.C., 1969).
3. J. M. Ziman, *Public Knowledge* (Cambridge University Press, New York, 1968).
4. Committee on Scientific and Technical Communication, *Report of the Task Group on the Economics of Primary Publication* (National Academy of Sciences, Washington, D.C., 1970), p. 8.
5. N. B. Hannay, *Cost Effectiveness of Information Systems. A Report by the Subcommittee on the Economics of Chemical Information of the Committee on Corporate Associates* (American Chemical Society, Washington, D.C., 1969).
6. J. Bernard, C. Schilling, and J. Tyson, *Informal Communication Among Bioscientists* (Biological Sciences Communication Project, George Washington University, Washington, D.C., 1963).
7. C. Schilling and J. Bernard, *Informal Communication Among Bioscientists, Part II* (Biological Sciences Communication Project, George Washington University, Washington, D.C., 1964).
8. F. J. Brockis and P. F. Cole, Chem. Brit. *3*, 421 (1967).
9. D. G. Marquis and T. J. Allen, Am. Psychol. *21*, 1052 (1966).
10. D. Marquis and D. M. Straight, "Organizational Factors in Project Performance," M. C. Yovits *et al.*, *Research Program Effectiveness* (Gordon & Beach, New York, 1966).
11. M. Slater and P. Fisher, *Use Made of Technical Libraries* (ASLIB Research Department, London, 1969).
12. W. M. Carlson, Trans. N.Y. Acad. Sci. *31*, 803 (1969).
13. M. H. Halbert and R. L. Ackoff, in *Proceedings of the International Conference on Scientific Information* (National Academy of Sciences–National Research Council, Washington, D.C., 1959), p. 970.
14. Case Institute of Technology, Operations Research Group, *An Operational Research Study of the Dissemination and Use of Recorded Scientific Information: A Study under NSF G-8434* (Report PB 171503 available from the National Technical Information Service).
15. M. W. Martin, Jr., IRE Trans. Eng. Management *EM-6*, 66 (1962).
16. M. W. Martin and R. L. Ackoff, Management Sci. *9*, 322 (1953).
17. S. Herner, in *Proceedings of the International Conference on Scientific Information* (National Academy of Sciences–National Research Council, Washington, D.C., 1959), p. 277.
18. M. A. Libbey and G. Zaltman, *The Role and Distribution of Written Informal Communication in Theoretical High Energy Physics*, Report AIP/SDD-1 (American Institute of Physics, New York, 1967) or USAEC Report No. NYO-3732-1.
19. R. M. Fishenden, in *Proceedings of the International Conference on Scientific Information* (National Academy of Sciences–National Research Council, Washington, D.C., 1959), p. 163.
20. B. Glass and S. Norwood, in *Proceedings of the International Conference on*

Scientific Information (National Academy of Sciences–National Research Council, Washington, D.C., 1959), p. 195.

21. M. Slater and S. Keenan, *Results of Questionnaire on Current Awareness Used by Physicists Prior to Publication of "Current Papers in Physics,"* Report AIP/CPP-1 (American Institute of Physics, New York, 1967). A partial summary is given in Keenan and Slater.[52]
22. Columbia University, Bureau of Applied Social Research, *Formal and Informal Satisfaction of the Information Requirements of Chemists,* Interim Report, H. Menzel, Project Director (New York, 1966).
23. J. Martyn, *Report of an Investigation in Literature Searching by Research Scientists* (ASLIB Research Department, London, 1964); New Scientist *21,* 338.
24. V. Rosenberg, Inf. Storage Retrieval *3,* 119 (1967).
25. P. G. Gerstberger and T. J. Allen, J. Appl. Psychol. *52,* 272 (1968).
26. T. J. Allen, in *Communication among Scientists and Engineers,* C. E. Nelson and D. K. Pollock, eds. (Heath Lexington Books, Lexington, Mass., 1970), p. 191.
27. M. Kotani, "Communication among Japanese Scientists Domestically and with Their Counterparts Abroad" (10th Pacific Congress, Honolulu, Hawaii, 1951).
28. Project on Scientific Information Exchange in Psychology, *Report of the American Psychological Association on Scientific Information Exchange in Psychology* (American Psychological Association, Washington, D.C., 1963), Vol. 1.
29. Project on Scientific Information Exchange in Psychology, *Report of the American Psychological Association on Scientific Information Exchange in Psychology* (American Psychological Association, Washington, D.C., 1965), Vol. 2.
30. B. C. Vickery, *Techniques of Information Retrieval* (Butterworths, London, 1970); see especially Chapter V Appendix.
31. W. O. Hagstrom, in *Communication among Scientists and Engineers,* C. E. Nelson and D. K. Pollock, eds. (Heath Lexington Books, Lexington, Mass., 1970), p. 85.
32. R. von Zelst and W. A. Kerr, J. Abnormal Soc. Psychol. *46,* 470 (1951).
33. R. E. Maizell, Am. Doc. *11,* 9 (1960).
34. E. T. Klemmer and F. W. Snyder, J. Commun. *22,* 142 (1972).
35. J. R. Hinrichs, Personnel Psychol. *17,* 193 (1964).
36. T. J. Allen, "Managing the Flow of Scientific and Technical Information" (Ph.D. thesis, MIT, Cambridge, Mass., 1966).
37. T. J. Allen and S. I. Cohen, Admin. Sci. Quart. *14,* 12 (1969).
38. L. J. Anthony, H. East, and M. J. Slater, Rep. Progr. Phys. *32,* 709 (1969).
39. S. Keenan, Library Trends *16,* 329 (1968); also available as AIP Report ID-687.
40. R. H. Whitford, *Physics Literature,* 2nd ed. (Scarecrow Press, Metuchen, N.J., 1968).
41. M. Terry and E. Cooper, *Secondary Services in Physics,* Report ID 69-2 (American Institute of Physics, New York, 1969).
42. American Institute of Physics, *Searchable Physics Information Notices, Technical Specifications,* Report ID 71-S (AIP, New York, 1971).
43. J. H. Connor, Library Quart. *37,* 373 (1967).
44. H. East, ASLIB Proc. *20,* 482 (1968).
45. International Council of Scientific Unions, Abstracting Board, *World Literature in Physics as Seen through Bulletin Signaletique 1964 Issues* (ICSU, Paris, France, 1967).

46. International Council of Scientific Unions, Abstracting Board, *World Literature in Physics as Seen through Physics Abstracts 1964 Issues* (ICSU, Paris, France, 1967).
47. R. C. Greer and F. Atherton, *Study of Nuclear Science Abstracts and Physics Abstracts Coverage of Physics Journals 1964–65*, Report AIP/DRP 66-1 (American Institute of Physics, New York, 1966).
48. C. Herring, Part I, in *Perspectives in Materials Research*, L. Himmel, J. J. Harwood, and W. J. Harris, Jr., eds. (Office of Naval Research, Washington, D.C., 1961), especially Figure 1 and Table 1.
49. S. Keenan and F. G. Brickwedde, *Journal Literature Covered by Physics Abstracts in 1965*, Report ID 68-1 (American Institute of Physics, New York, 1968).
50. E. J. Feinler, C. J. Cook, O. Heinz, and C. P. Bourne, Am. Doc. *16*, 329 (1965).
51. P. Atherton, *A Preliminary Report on Phase I of the Reference Retrieval System Development Project*, Report AIP/DRP 62-2 (American Institute of Physics, New York, 1962).
52. S. Keenan and M. Slater, J. Doc. *24*, 98 (1968).
53. M. Slater and S. Keenan, *Current Papers in Physics User Study: Coverage, Arrangement, and Format*, Report AIP/CPP 2 (American Institute of Physics, New York, 1968).
54. M. Slater and S. Keenan, *Use Made of Current Papers in Physics*, Report AIP/CPP 3 (American Institute of Physics, New York, 1968).
55. UNISIST, *Study Report on the Feasibility of a World Science Information System by the United Nations Educational, Scientific and Cultural Organization and the International Council of Scientific Unions* (UNESCO, Paris, 1971).
56. A. Herschman, Phys. Today *24*(11), 23 (1971).
57. N. Price and S. Schiminovich, Inf. Storage Retrieval *4*, 271 (1968).
58. S. Schiminovich, Inf. Storage Retrieval *6*, 417 (1971).
59. K. D. Carroll, *Survey of Scientific-Technical Tape Services*, Report AIP/ID 70-3 (American Institute of Physics, New York, 1970); or Report ASIS SIG/SDI 2 (American Society for Information Science, Washington, D.C.).
60. E. B. Parker, in *Communication among Scientists and Engineers*, C. E. Nelson and D. K. Pollock, eds. (Heath Lexington Books, Lexington, Mass., 1970), p. 281.
61. M. M. Kessler, Phys. Today *18*(3), 28 (1965).
62. *Project INTREX* (Semiannual Activity Reports) (Massachusetts Institute of Technology, Cambridge, Mass., 1966–1971).
63. J. Vlachy, Czech. J. Phys. *B20*, 501 (1970).
64. S. Keenan and P. Atherton, *The Journal Literature of Physics*, Report AIP/DRP 9A1 (American Institute of Physics, New York, 1966).
65. International Council of Scientific Unions, Abstracting Board, *Some Characteristics of the Primary Journals in the Domain of the Physical Sciences* (ICSU, Paris, France, 1966).
66. M. Cooper and C. W. Thayer, *Primary Journal Literature of Physics*, Report ID 69-4 (American Institute of Physics, New York, 1969).
67. M. B. Gardner, J. Audio Eng. Soc. *19*, 560 (1971).
68. H. W. Koch, *Economics of Primary Journals in Physics*, Report ID 69-5 (American Institute of Physics, New York, 1969).
69. H. A. Barton, Phys. Today *16*(6), 45 (1963).
70. Federal Council for Science and Technology, Sci. Inf. Notes *3*(5), 1 (1961).

71. M. A. Tinker, *Legibility of Print* (Iowa State University Press, Ames, 1963).
72. H. Zuckerman and R. K. Merton, Minerva *9,* 66 (1971).
73. H. Zuckerman and R. K. Merton, Phys. Today *24*(7), 28 (1971).
74. H. W. Koch, *The Role of the Primary Journal in Physics,* Report ID 70-1 (American Institute of Physics, New York, 1970).
75. National Science Foundation, *Federal Funds for Research, Development, and Other Scientific Activities, Fiscal Years 1968, 1969, and 1970,* Document NSF 69-31 (Washington, D.C., 1969).
76. J. Vlachy, Teorie a Metoda *3,* 83 (1971).
77. R. E. Burton and R. W. Kebler, Am. Doc. *11,* 18 (1960).
78. D. J. de S. Price, Science *149,* 510 (1965).
79. D. J. de S. Price, in *Communication among Scientists and Engineers,* C. E. Nelson and D. K. Pollock, eds. (Heath Lexington Books, Lexington, Mass., 1970), p. 3.
80. Federal Council for Science and Technology, Committee on Scientific and Technical Information, *Panel Report on the Role of the Technical Report in Scientific and Technical Communication* (Washington, D.C., 1967).
81. D. E. Gray and S. Rosenborg, Phys. Today *10*(6), 18 (1957).
82. M. J. Moravcsik, Phys. Today *18*(3), 23 (1965).
83. A Debate on Preprint Exchange (remarks by M. J. Moravcsik and S. Pasternack), Phys. Today *19*(6), 62, 63, 71 (1966).
84. D. Green, Int. Sci. Technol. p. 82 (May 1967).
85. P. Cole, J. Doc. *19,* 1 (1963).
86. P. M. Strain, Library Res. Technol. *10,* 295 (1966).
87. D. Gushee, J. Chem. Doc. *8,* 191 (1968).
88. W. D. Garvey and B. C. Griffith, Science *146,* 1655 (1964).
89. R. E. Burton and B. A. Green, Phys. Today *14*(10), 35 (1961).
90. S. A. Goudsmit, Phys. Today *19*(9), 52 (1966).
91. C. Herring, J. Chem. Doc. *8,* 232 (1968).
92. Ad Hoc Panel on Nuclear Science, National Research Council, *Nuclear Data Compilations: The Lifeblood of the Nuclear Sciences and Their Applications* (National Academy of Sciences, Washington, D.C., 1971).
93. J. D. Bernal, J. Doc. *13,* 195 (1957).
94. C. Fix, D. T. H. Campbell, and W. A. Creager, *Some Characteristics of the Review Literature in Eight Fields of Science,* Report to the Office of Science Information Service, National Science Foundation, 1964 (Herner & Company).
95. J. Stern, unpublished work at the American Institute of Physics (1969).
96. C. Herring, Phys. Today *21*(9), 27 (1968).
97. G. T. Rado and H. Suhl, eds., *Magnetism II B* (Academic Press, New York, 1966).
98. D. A. Brunning, "Review Literature and the Chemist," in *Proceedings of the International Conference on Scientific Information* (National Academy of Sciences–National Research Council, Washington, D.C., 1959), Vol. 1, p. 545.
100. American Institute of Physics, *The AIP Program for Physics Information* (AIP, New York, 1971).
101. Johns Hopkins University, *The Dissemination of Scientific Information, Informal Interaction, and the Impact of Information Received from Two Meetings of the Optical Society of America,* Report No. 3 of the Johns Hopkins University Center for Research in Scientific Communications (Baltimore, Md., 1970).
102. B. E. Compton, Technol. Soc. *4*(2), 39 (1968).

103. W. D. Garvey, in *Communication among Scientists and Engineers,* C. E. Nelson and R. K. Pollock, eds. (Heath Lexington Books, Lexington, Mass., 1970).
104. R. S. Rosenbloom and F. W. Wolek, *Technology, Information, and Organization/Information Transfer in Industrial R&D* (Harvard University, Graduate School of Business Administration, Cambridge, Mass., 1967).
105. S. Herner, Ind. Eng. Chem. *46,* 228 (1954).
106. Federal Council for Science and Technology, Committee on Scientific and Technical Information, *Directory of Federally Supported Information Analysis Centers* (Washington, D.C., 1968).
107. E. L. Brady, in *The Information Analysis Center–Seven Background Papers,* Report COSATI 69-6 (Federal Council for Science and Technology, Committee on Scientific and Technical Information, Washington, D.C., 1969).
108. National Bureau of Standards, *Critical Evaluation of Data in the Physical Sciences/A Status Report on the National Standard Reference Data System,* Technical Note 553 (Washington, D.C., 1970).
109. D. R. Stein, Spec. Libraries *50,* 492 (1959).
110. K. C. Buschbeck, W. Lippert, and E. Uhlein, Naturwiss. *55,* 379 (1968).
111. A. Herschman, in *Communication among Scientists and Engineers,* C. E. Nelson and D. K. Pollock, eds. (Heath Lexington Books, Lexington, Mass., 1970).
112. I. M. Orient, quoted in V. V. Nalimov and Z. M. Mul'chenko, Naukometriya (Izdatel'stvo "Nauka," Moscow, 1969), p. 71.
113. W. Lippert, Chem. Z. *94,* 47 (1970).
114. *A Program for a National Information System for Physics and Astronomy 1971–1975,* Report No. AIP/ID 70-P (Information Division, American Institute of Physics, New York, 1970).
115. C. Benjamin, Sci. Res. p. 32 (Sept. 16, 1968).

APPENDIX: PHYSICS SURVEY—A CHARGE TO THE SURVEY PANELS

The following are topics on which the Survey Committee requests input information from the Panels:

THE NATURE OF THE FIELD

It is vitally important that we communicate to our audiences some coherent presentation of what we believe physics is all about. Please help us in this by considering how best to present your field to (a) other physicists, (b) other scientists, (c) nonscientists. Particularly in the latter case it will be helpful to provide the Committee with what the Panel may well consider an oversimplified and overpopularized view—previous panels have erred in the opposite sense. Examples, illustrations, case history—and indeed some historical perspective generally—will be most helpful.

THE STATUS OF THE FIELD

(a) What have been the major developments (both in theory and experiment) during the past five years? If possible, put these into context with reference to the status statements in the Pake Survey and Panel reports.

(b) What are the implications of these developments for the growth of the field during the next five years?

(c) What are specific examples of major changes or advances that these new developments afford? Can we do things now that were simply impossible before? Are there examples that could provide striking graphic treatment in our report?

(d) What are the present frontier areas of the field? How are these defined?

(e) Is the balance between experimental and theoretical activity in the field at a desirable level? If not, what are the Panel recommendations concerning an optimum balance and how might it be achieved?

Institutions of the Field

(a) How is activity in this field now divided among the various types of research institutions, i.e., academic, national laboratory (e.g., Brookhaven), government laboratory (e.g., NRL), industrial laboratory, etc.?

(b) What recommendations does the Panel have concerning this balance and its possible modification in the next five years? The next decade?

(c) In this field, what are the characteristic features of activity in the different institutions?

(d) What are the interactions between these institutions? Are there areas where this interaction could or should be improved? What are the effective barriers, if any, that may prevent ready communication between, or direct exchange of, personnel for example?

Interaction with Other Areas of Physics

(a) Illustrating with specific examples wherever possible, what have been the outstanding examples of interaction between this and other fields of physics recognizing that this is almost always a two-way interaction?

(b) What are specific examples of techniques—either experimental or theoretical—that cut across field boundaries? Detailed studies of selected examples would be particularly useful.

Interaction with Other Areas of Science

Questions identical to those above seem appropriate again with stress on the desirability of specific examples and possible illustrative material. The most important interactions will, of course, vary with the field; areas such as chemistry, medicine, biological sciences, ecology are obvious candidates for consideration.

Interaction with Technology

Research and technology have long advanced through mutual stimulation. In this field, what are the outstanding examples of such interaction in recent years? Case studies are particularly useful here. Purely as an example that has been suggested, it might be useful to consider an essay covering a tour through a modern hospital, a chemical processing plant, a paper mill, or the like, noting in passing those techniques and instruments that have arisen from work in the field. Cooperative efforts with other Panels would seem profitable. The inverse should not be neglected; some emphasis on the great dependence of research progress on technological progress is clearly indicated.

The Data Panel will attempt to arrive at methods of quantifying some of the available information in the area—both within and outside of this country. Close collaboration with the Data Panel in identifying areas of particular importance and interest would be most helpful.

Interaction with Industry

(a) Illustrating, wherever possible, with *specific* examples, what have been the outstanding interactions between this field of physics and the industrial sector in the past five years?

(b) Can any of the recent developments in the field be extrapolated, at this time, as having such interaction in the near, or distant, future?

(c) What is the inverse situation? What impact have techniques, products, or people in the industrial sector had on this field?

(d) How can the interaction between this field and the industrial sector be made more effective?

(e) It has been suggested that the development of biotechnology represents the conversion of the last of the guilds into an industry. What contributions has this field made to this conversion?

(f) Succinct case studies would be very valuable here.

Interaction with Society

(a) In what areas is the field already having major impact on questions of direct social importance?

(b) What other areas are candidates for such interaction?

(c) What aspects of training in this field are of particular importance for utilization in problems of broader social implication—which of these latter in particular?

(d) What would be the Panel recommendations concerning broader utilization of present personnel and facilities on such problems? Examples of possible situations would be most helpful.

(e) A few groups have already decided to devote some selected fraction of their effort to such activities. A discussion of such approaches would be helpful.

(f) One of the major questions facing physics (and science generally) is that of educating the nonscientific public to its very real relevance—however defined—in a technological civilization. The Survey Committee would welcome suggestions, case histories, examples, and any material that would assist in its consideration of this question for physics generally, as well as more specifically within the context of the Panel's subfield.

Relationship to Other Areas of Social Concern

Traditionally, physics has been recognized as being relevant to national defense, atomic energy, space, etc., and has enjoyed support from the corresponding federal agencies. Today our society is moving its center of concern to areas for which, at first sight, physics is less relevant: health, pollution, racial tension, etc. The new federal agencies organized to deal with such questions, such as NIH, HUD, DOT, accept much less, or no, responsibility for physics. How strong a case can be made for the relevance of your subfield to the achievement of the missions of these other agencies? In general, this will come through the help your field can give to technologies that will further these social ends: for example, the role of computers (and therefore solid-state physics) in automating hospital care. However, there may be other more direct inputs to your field that do not go through technology.

Cultural Aspects of Physics

Knowledge of the physical universe has more than utilitarian value. Each advance in fundamental understanding becomes an indestructible asset of all educated men. It is not suggested that each Panel should provide an essay on the contributions of its field to human culture, but it would be helpful in developing a broad exposition of this aspect of physics to have suggestions or compelling examples related to your field. A rather obvious concrete example: we know how old the earth is; that knowledge came through physics. Examples less obvious, and especially examples of important questions that may be answered in the foreseeable future, would be welcome.

We would welcome assistance from the Panel in answering such questions as (a) How best do we bring out the cultural relevance of physics?

(b) To what extent should our report develop the cultural arguments as a basic justification for continuing support of physics? (c) How can we best address ourselves to the resurgence of mysticism and of anti-intellectual and antiscience attitudes among students? Among the citizenry generally? (d) What is the role of physics in countering these developments?

Relationship to National Security Activities

(a) What role has the field played in national defense activities?

(b) What future role is envisaged? How important is the field to these activities? Disarmament activities should be carefully considered in this context.

(c) What have been the respective roles of the different institutions of physics in this area?

(d) Again the Committee would welcome the assistance of the Panel in addressing the general questions relating to the overall interaction of physics in national security activities.

Training in the Field

(a) It is often implied that contemporary graduate and postdoctoral training is becoming so narrow that students have lost the traditional breadth of outlook and flexibility expected of a physicist. Is this situation true in this field? What can be done to improve the situation? What recommendations does the Panel have for modification of contemporary training programs?

(b) In what ways is this field of particular importance for physics education?

(c) Although clearly the question relates to all of physics, can the Panel provide relevant input to the Committee concerning (i) the adequacy of current secondary school training in physics and mathematics; (ii) the effectiveness of some of the more modern secondary school curricula, e.g., PSSC; (iii) the relative intellectual standing, at the secondary school level, of those students who choose to major in undergraduate physics? (There is a widespread element of folklore that suggests that physics no longer attracts the most intellectually gifted secondary students. Can this be supported or refuted? What is the significance of this statistic in whichever case emerges?)

(d) Again, although relating to all of physics rather than to this Panel specifically, the Committee would welcome input concerning such topics as (i) what has been accomplished in bridging the gap between physics and other disciplines at the undergraduate level? How successful have general science or interdisciplinary courses been for entering—for advanced—students? How can we better illustrate the fundamental impact of physics as an underlying discipline in many areas of undergraduate education? (ii) Are teaching materials adequate? Do presently used textbooks adequately reflect the contemporary structure of physics? (iii) How important a demand for trained physicists will teaching requirements represent at established university centers—at newer campuses—at the colleges?

(e) To what extent has obsolescence of training overtaken members of the field? What can be done about it?

(f) What effective mid-career training opportunities now exist in the field? What are the Panel recommendations in this area?

(g) How effective are existing summer school programs in meeting the need for continuing training and education in the field?

(h) How effective are conferences and symposia in the field as training mechanisms?

(i) What are the Panel recommendations concerning the number and character of such conferences and symposia now available in the field?

Postdoctoral Training

(a) What is the role of the postdoctoral appointment in the field? This will, of course, be different in the different institutions.

(b) What is the average duration of the postdoctoral appointment? How has this changed with time?

(c) What has been the distribution, by nationality, of postdoctoral people in the field, and what fraction of these have remained in the United States following their postdoctoral training? How has this changed with time?

(d) How has the leveling of funding affected the availability of postdoctoral appointments in the different institutions (e.g., industrial laboratories, national laboratories, government laboratories, universities)?

Training in Applied Areas of the Field

(a) What are the applied areas that draw most heavily on this field?

(b) Does the supply of physicists in this field suffice to meet the demand in these areas?

(c) Is the current training adequate? Would modification of current training patterns be expected to open up significant new employment opportunities?

(d) It might be argued that there has been a significant failure in communication between prospective applied physics employers and the academic groups involved in the applied training. Is this true in this field? If so, how can it be improved?

(e) How is the applied work distributed with regard to the type of institutions involved?

Manpower Projections

(a) What is the current population in the field, and how has this population developed since 1965 (as covered in the Pake reports) in (i) academic research, (ii) industrial research, (iii) government laboratory research, (iv) postdoctoral training, (v) graduate student training, (vi) other?

(b) During the same period what migration has occurred into—and out of—the field? What have been the major sources and recipients of this migration?

(c) In the light of current challenges in the field and/or new or anticipated facilities, what projected manpower needs can be expected in each of the above area in the next five years—the next ten (recognizing that this latter is an extreme extrapolation at best and closely related to available funding)?

(d) The argument is often advanced that the shortage of jobs requires additional funding in the field. This is more frequently reversed in Washington to imply simply that there are too many physicists being trained. What is the situation in this field?

(e) To what extent is the claim of inadequate employment opportunities legitimate (i.e., to what extent does this simply reflect the fact that for perhaps the first time physicists are not able to obtain the job that they would find most attractive)? What fraction of current PhD graduates were unsuccessful in finding employment where they were in a position to utilize their broad physics training if not their immediate specialty training?

(f) Will adequate manpower be available to staff emerging institutions in the field? How can qualified staff be attracted to and retained by such institutions?

(g) Does this field have unique or special characteristics that recommend it for consideration by an emerging institution?

(h) With leveling funding it may well be impossible for new (and indeed old) institutions to span as broad a spectrum of fields of physics as has been traditional, and while regrettable from a training viewpoint further specialization may be required in any given institution. How feasible are joint activities in this field as compared to others in physics? What recommendations would the Panel have in this difficult area?

Facilities

(a) Existing Facilities

(i) What are the major facilities in the field, and how are they distributed as to type?

(ii) Are the existing facilities now being utilized to full capacity? If not, explain.

(iii) How are present facilities being utilized, i.e., are they shared by more than a single group, how are decisions made regarding the research scheduling?

(iv) What are the outstanding problems now faced in the use of existing facilities?

(v) Is the distribution of existing facilities adequate?

(vi) Is modernization of the existing facilities feasible? What is the estimated effective lifetime of typical existing facilities in the field?

(vii) What criteria should be applied in reaching decisions to close down existing facilities?

(viii) To what extent do such criteria differ in different institutions (e.g., a facility might have training potential in an academic environment when it has reached a stage of unacceptable obsolescence elsewhere)? Is relocation of facilities a viable suggestion under such conditions? There are clearly pitfalls of which the receiving institution should be aware. What are they in this field?

(b) New Facilities

(i) What new facilities will be required to exploit the potential of the field? What is the priority ordering of these facilities? Please support with detailed discussion.

(ii) To what extent could existing facilities now used by other areas of physics be adapted for frontier use in this field?

(iii) What are the Panel recommendations regarding siting and operation of new facilities?

(iv) Within this field, what is an optimum balance between large centralized facilities and smaller more widely distributed ones? Please discuss.

(v) What new developments now on the horizon show promise of evolution as major facilities in the field? Is an estimate of the probable gestation period and possible cost now possible for each?

The Impact of Computer Techniques on the Field

(a) What have been the outstanding impacts of computer technology in this field?

(b) Would larger and/or faster computers be of significant value? What would be the relative priority assigned to the higher costs that would be involved here as compared to other major capital needs of the field?

(c) Has any particular scheme of utilization, i.e., small local computers, institutional computer centers, regional computer centers emerged as preferable in this field?

(d) Do existing software and languages pose significant limitations in the field?

(e) What estimate does the Panel have for the present and projected utilization of computers in the field? Can a dollar level be attached to this?

(f) What impact has the field had on computer technology?

(g) Are there outstanding examples of studies that would simply have been impossible without sophisticated computer utilization? Specific examples would be most useful.

COST INCREASES

(a) Selecting, say, ten instruments much used in the field spanning the cost range involved—how have the individual costs varied with time in the last decade?

(b) How has the average (very crudely defined) overall cost of an experiment, typical of those at the frontier of the field at the time, varied with time in the last decade?

(c) How have average postdoctoral and student training costs varied over the same interval? It would be advantageous to consider experimental and theoretical situations separately in this instance.

(d) Illustrating with specific examples, what would be a reasonable annual estimate of the cost escalation in the field reflecting increasing sophistication of the studies themselves? Reflecting aging of the institutional staff?

(e) To what extent is progress in the field *really* dependent upon the availability of the most modern instrumentation? It has been suggested that in some fields the instrumentation has become over-sophisticated, overflossy and that in at least some instances the Ferrari could be replaced by a Ford without undue restriction of the research quality and productivity. To what extent is this suggestion true in this field? To what extent can (and should) it be countered? Specific illustrations and examples would be extremely helpful here.

FUNDING LEVELS

(a) What have been the actual funding levels *and* expenditure levels annually in the field since 1965? Compare these with the Pake Report projections. Insofar as possible, separate academic, industrial, and governmental laboratory operations for consideration. In some instances the leveling off of federal funding has been counteracted, for a time at least, by infusions of institutional funds, so that actual expenditure levels have not tracked funding limitations. What information is available on such phenomena in this field?

Funding Mechanisms

(a) How has the available funding been distributed among these sources: federal (AEC, DOD, NSF, NASA, others), state, industrial, local (university contributions, etc.), foundations, and other sources?

(b) How does the funding process actually work for each of the above sources? What are the relative distributions, advantages, disadvantages, etc., of grants and of contracts? What are the effective differences between these two approaches? What improvements might be suggested?

(c) What is the relative importance of project and of institutional grants in this field?

(d) How are decisions reached concerning grant and contract applications? Please comment on the decision-making processes at the national level—for example, by administrators in the various federal agencies and by advisory committees to these agencies. Is the present practice satisfactory or would change be desirable? What are the Panel recommendations?

The Impact of Leveling Funding

(a) Discuss in some detail, with specific illustrations, the overall impact of leveling funding on the field. The following subtopics might prove useful:

(i) Utilization of current facilities

(ii) Exploitation of new discoveries

(iii) Employment of physicists

(iv) Support of the young researcher

(v) Alienation of young physicists

(vi) Possible new approaches to training in the field

(vii) The support of offbeat proposals. There is always a tendency, under limited funding conditions, to eschew risk or adventure, to bet on the sure thing.

(viii) Long-range implications for the field generally.

(b) It is clear that level funding is not synonymous with level productivity. The Committee will welcome case histories, etc., to illustrate this general point.

(c) What are the relative advantages of expanding (or contracting) activities in this field by expanding (or contracting) the size of existing groups active in the field as opposed to proliferating (or reducing) the number of such groups?

FUNDING PROJECTIONS

In the past, survey reports have generally made specific projections and recommendations which have very often been negated by large departures of the total budgets available from those on which the recommendations were based. To be responsive, our Report must provide for a spectrum of possible situations; in doing so it must carefully spell out, in as detailed fashion as possible, both the short- and long-range consequences of funding at levels below those necessary for both orderly growth and exploitation of new developments in each of the fields of physics. Specific examples and case histories will be particularly effective in illustrating these consequences. With these points in mind,

(a) What level of funding, quite apart from any current estimate of future funding, would be required to enable this field to realize its full potential during the next five years? The next ten years? How would it be distributed broadly over the subareas of the field—recognizing that detailed projections are, in many cases, impossible?

(b) Consider a spectrum of possibilities ranging downward in 10% increments from that developed above to a level some 10% below that currently in effect. At each step indicate as clearly as possible,

(i) What opportunities would be missed—what developments would not be exploited?

(ii) What new facilities would necessarily be postponed or eliminated entirely from consideration?

(iii) What programs or facilities would necessarily be phased out or closed down?

(iv) What would the impact be on the manpower and employment situation?

(c) A detailed discussion of the basic issues that underlie the Panel's assignment of priorities within the field would be an essential component of the Panel report. It is essential that long-range implications be developed realistically; it is essential that we not predict greater catastrophic impact than can be clearly justified.

(d) Separate discussion of major new facilities—in order of priority—with careful discussion of the bases for the priority ordering and of the relative justifications will be particularly important.

(e) The question of laboratories, as distinct from facilities, will be appropriate in some fields. The need for and justification of such laboratories will require careful consideration. What are the recommended criteria for closing down an existing laboratory in this field? To what extent are the laboratories in the field adaptable to broader use and to alternate modes of support during periods of fiscal stringency?

(f) To what extent can the Panel assist in developing a balanced presentation of the overall impact on the continuity of physics (i.e., the faucet effect—it is not generally appreciated that the re-emergence of funding after an indeterminate drought will not guarantee re-emergence of a healthy physics—or science—community)? Can this be quantified in this field? Are there relevant examples or case histories?

(g) A clear statement of the basic fiscal assumptions underlying the Panel projections is essential. The Data Panel will provide basic information concerning inflation rates, etc., which should be used systematically by all Panels to permit later direct comparisons by the Committee.

Physics Data in the Field

(a) How effective is communication of scientific information in the field generally? Are there adequate review articles—conferences and conference proceedings? Are there too many of the latter?

(b) What is the role of the preprint in this field? Is the present system effective?

(c) How adequate are the present data compilation and dissemination mechanisms in this field?

(d) What are the Panel recommendations in this area? Are new approaches or mechanisms required? How can manpower, adequate both in quantity and quality, be integrated into the data compilation activities?

(e) What is the estimated cost involved?

(f) Quite apart from data communication and compilation within the field, (i) how effective is communication with related fields that may have need of your data, and (ii) how effective are your data formats and presentations for their use?

International Aspects

(a) Where does this field in the United States at the present time stand with respect to the same field abroad?

(b) How does U.S. activity in the field compare on a manpower or funding basis with that in the most active foreign countries? What are the relative growth rates? What are the major points of similarity or dissimilarity in the overall programs? What has been the significance of the different funding techniques and levels?

(c) What international cooperation now exists? What would be the direct and indirect benefits to the United States in expanding such cooperation in this field? Are there particular facilities that should be considered in this light?

(d) What problems now exist with regard to the implementation of foreign cooperation and exchanges? Have problems been encountered in the obtaining of requisite visas—of permission to travel freely across international boundaries—of access to national or governmental laboratories in this country or abroad?

(e) What is the situation vis-à-vis international cooperation in physics in the industrial sector? Are there outstanding difficulties in this area? How important is fostering of such cooperation in this field?

(f) To what extent does this field encompass well-defined national schools of thought (e.g., the Copenhagen School in quantum mechanics and nuclear physics)?

(g) What has been the impact of foreign work and foreign research centers on activity in this field in this country?

(h) How do developing countries attain critical mass in this field? Are there specific mechanisms in this area? Should there be?

(i) What international laboratories *should* be developed in this field? Upon what criteria should the establishment of such laboratories be based?

Illustrative Material for the Survey Report

It will be particularly important that the Committee receive from each Panel a selection of illustrations and photographs carefully selected to highlight progress or particularly interesting vignettes in each field. It would be helpful if the Panels would address themselves to this request at an early stage of their deliberations. The members of the Data Panel will devote considerable effort to the development of new techniques for the presentation of statistical data and will cooperate closely with each of the subfield Panels.