NASA SP-5078

TECHNOLOGY UTILIZATION

BIOMEDICAL RESEARCH AND COMPUTER APPLICATION IN MANNED SPACE FLIGHT

A REPORT

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

For sale by the Superintendent of Documents,
U.S. Government Printing Office, Washington, D. C. 20402
Price $2
Library of Congress Catalog Card Number 72–608485

FOREWORD

The National Aeronautics and Space Administration has established a Technology Utilization Program designed to transfer technological developments that may have useful commercial applications. From NASA laboratories and contractors, aeronautics and space-related technology is gathered and evaluated. Items which have potential industrial use are made generally available. This survey of computer uses in the field of medicine is one of a series of NASA publications that presents information of direct or indirect interest to the nonaerospace community.

During the past twenty-five years, computers have grown from slow-speed vacuum-tube machines with very small storage capacity and very limited processing capability to very fast, very large, extremely versatile devices that use microminiature electronic components. During the same period of time, there has been an exponential growth of knowledge in the medical sciences. To a considerable degree, America's twelve-year old program of space exploration has stimulated development in both areas.

With the increase of medical knowledge has come a multiplication in the quantity and kinds of medical data. It seems logical to expect that the computer would be applied to aid in the reduction and interpretation of this data, but the full potential of computers in medicine has not yet been realized.

This report summarizes the areas of medicine in which computers can be employed and examines in detail several cases where computers have been applied in connection with the medical aspects of NASA's manned space flight program. Treated are such problems as those of automated medical data storage and retrieval systems, continuous monitoring and interpretation of electrocardiograms, and computer-aided medical diagnosis. The approach is cautious throughout, with the emphasis almost constantly on ways to permit the computer to perform various clerical functions while leaving critical decisions to a human monitor.

RONALD J. PHILIPS, Director
Technology Utilization Office

Page intentionally left blank

PREFACE

Research into the biomedical variables surrounding safe and efficient manned space flight has revealed the need for the advanced application of computers in the solution of various problems. This book presents manned spaceflight-oriented research that emphasizes the computerized approach both in terms of a general overview and as utilized in the solution of specific problem areas.

Since the articles presented here have not previously been made available to the general scientific community, it is believed that this book will preserve and disseminate much knowledge concerning manned space flight research and the general and specifically related application of computers.

This organized collection of articles resulted from activities encouraged, and in most cases supported, by the National Aeronautics and Space Administration. The material presented here spans the tenures of George M. Knauf, M.D.; William Randolph Lovelace, II, M.D.; and Brigadier General Jack Bollerud, USAF, MC, in the Office of the Director of Space Medicine, NASA Headquarters. Concurrently, under each of the foregoing Directors, Dr. Jefferson F. Lindsey, Jr., served as Head of the Biomathematical Staff.

The efforts of Charles A. Berry, M.D., and A. Duane Catterson, M.D., of the Manned Spacecraft Center, Houston, Texas, in furthering the research efforts reported here are to be particularly noted. Likewise, appreciation is expressed to Mr. Walter B. Sullivan, Jr., NASA Headquarters, who was directly concerned as monitor on several of these projects, and to Dr. Mae M. Link, NASA Headquarters, for her coordination efforts in expediting the publication of this book.

J. W. Humphreys, Jr.
Major General, USAF, MC
Director, Space Medicine
Manned Space Flight

CONTRIBUTORS AND EDITORS

Abraham, Sidney: Statistician, Medical Systems Development Laboratory, Heart Disease Control Program, National Center for Chronic Disease Control, Public Health Service, U.S. Department of Health, Education, and Welfare, Washington, D.C.

Burch, Neil R., M.D.: Head, Psychophysiology Division, Baylor University College of Medicine, Texas Medical Center, Houston, Texas

Caceres, Cesar A., M.D.: Chief, Instrumentation Field Station, Heart Disease Control Program; and Associate Professor of Medicine, George Washington University, Washington, D.C.

Calatayud, Juan B., M.D.: Assistant Professor of Medicine, George Washington University, Washington, D.C.

Dorsett, Ronald G.: Texas Research Institute of Mental Sciences, Baylor University College of Medicine, Texas Medical Center, Houston, Texas

Hochberg, Howard M., M.D.: Medical Officer, Medical Systems Development Laboratory

Horton, Caroline L., B.S.: Senior Computer Programmer, Section of Medical Information Management Systems, Department of Biomathematics, The University of Texas, M.D. Anderson Hospital and Tumor Institute, Houston, Texas

Jackson, Larry K., M.D.: Staff Physician, Medical Systems Development Laboratory

Jones, Robert L., Ph.D.: Associate Professor, Baylor University College of Medicine; formerly with the Biomedical Research Office, National Aeronautics and Space Administration, Manned Spacecraft Center, Houston, Texas

Knoblock, Edward C., Ph.D., Colonel, USA: Director, Division of Biochemistry, Walter Reed Army Institute of Research, Washington, D.C.

Lester, Boyd K., M.D.: University of Oklahoma Medical Center, Oklahoma City, Oklahoma

Lindsey, Jefferson F., Jr., Ed.D.: Assistant to the President, Southern Illinois University, Carbondale; formerly with the Office of Space Medicine, National Aeronautics and Space Administration, Washington, D.C.

McAllister, James W.: Engineer, Medical Systems Development Laboratory

Minckler, Tate M., M.D.: Chief, Section of Medical Information Management Systems, Department of Biomathematics, The University of Texas, M. D. Anderson Hospital and Tumor Institute, Houston, Texas

Moseley, Edward C., Ph.D.: Biomedical Research Office, National Aeronautics and Space Administration, Manned Spacecraft Center, Houston, Texas

Paine, Stephen T., Ph.D.: Life Sciences Operation, North American Aviation, Inc., El Segundo, California

Proctor, Lorne D., M.D.: Chairman, Department of Neurology and Psychiatry, Henry Ford Hospital, Detroit, Michigan

Rosner, Stuart W., M.D.: Medical Officer, Medical Systems Development Laboratory

Sashin, Jerome, M.D.: Medical Officer, Medical Systems Development Laboratory

Sells, S. B., Ph.D.: Research Professor (Psychology) and Director, Institute of Behavioral Research, Texas Christian University, Fort Worth, Texas

Stein, Mervyn R., M.D.: Clinical Laboratory Service, Clinical Center, National Institutes of Health, Bethesda, Maryland

Surgent, Louis V., Ph.D.: Research Staff Specialist, Research Department, Electronics Division, General Dynamics, Inc., Rochester, New York

Townsend, John C., Ph.D.: Professor and Director of the Human Factors Program, Department of Psychology, The Catholic University of America, Washington, D.C.; and Consultant to the Office of Space Medicine

Vordeman, Abbie L., M.D.: Texas Research Institute of Mental Sciences, Baylor University, College of Medicine, Texas Medical Center, Houston, Texas

Weihrer, Anna Lea: Mathematician, Medical Systems Development Laboratory

Winer, David E.: Engineer, Medical Systems Development Laboratory

CONTENTS

CHAPTER 1

MEDICAL USAGE OF COMPUTER SCIENCE

John C. Townsend

Since the early 1940's, the development of computers has been exponential. Unfortunately their utilization in certain fields of endeavor has lagged far behind their availability. The medical field affords an outstanding example of this lag, one reason apparently being the unfamiliarity of medical personnel with the availability of specific computer applications to the medical field. Because there is a growing body of literature relating computers to medicine, which could close the gap, reporting portions of the literature will contribute heavily to this chapter.

Certain observations made during collection of material for this chapter succinctly point up the status of computer utilization in the medical field. The observations were based on information gathered from 84 medical schools, many civilian and military installations, and recognized experts; the biomedical journals were surveyed also. In too many cases the working medical respondents stated that they were just then looking into computers and hoped to be able to use them in the near future. Often the apparent local interest in computer use was mostly held by nonmedical personnel working in medical settings. On the whole the response presented a picture of extremely spotty use by medical departments. Often where computers were available, and a list of the uses of computers by resident medical personnel was provided, there appeared, by inference from examination of the scope and methods of usages, to be lack of knowledge of the full potential of computers in a particular area of medicine. On the other hand, computers were being used very efficiently at some institutions.

The approach in this chapter will be separate treatment of the various categories of use of computers in the medical field. Although some categories overlap, a good deal of compartmentalization was not too difficult.

UTILIZATION OF COMPUTERS IN MEDICAL DIAGNOSIS

Since 1959 there have been several concerted efforts in research aimed at use of digital computers for medical diagnoses. The approach has been cautious. Even the most recent literature is peppered with such phrases as "of course the computers will only serve as an aid to the physician when he makes his diagnosis," and "the computer is to be used to make only presumptive diagnoses." However, the results of certain approaches seem to permit a more optimistic viewpoint. Let us look at the completed research and see just what can and what cannot be done in the area of medical diagnosis.

Use of the computer for medical diagnoses is really an attempt at simulation of the reasoning process and the memory of the physician in dealing with the patient's symptoms. The input data to the system are the symptoms. The literature reports several ways in which symptoms have been gathered from patients, as well as several ways in which the appropriate cluster of symptoms has been established for a particular disease.

Obtaining Input Data on Symptoms

Most commonly used in this regard is the self-administered printed form of the Cornell Medical Index – Health Questionnaire (CMI); it solicits yes-or-no answers to 195 questions covering the patient's medical history and complaint. The patient's age and sex are added as input information. The yes-no type of answer mates well with the binary input requirements of the computer. It is obvious, however, that the input of system information is mainly subjective.

In an attempt to evaluate the use of subjective symptoms in diagnosis of disease by computer, Rinaldo et al. (ref. 1) chose epigastric pain because it embodied the classic problems in symptom diagnosis. On the first interview, the patient was asked to answer "yes" or "no" to the following eight questions: Right-upper-quadrant pain? Clusters? Brief, irregular? Food relief? Food aggravation? Positional aggravation? Weight loss, 20 lb? Persistent? The patient's age also was obtained. These symptoms were solicited as predictors of six diagnostic categories: hiatus hernia, duodenal ulcer, gastric ulcer, gallstones, functional, and cancer. The six diagnoses were used because they were most comprehensive and made a list short enough for handling by the IBM-650 computing facilities.

Other investigators have dealt with more objective symptom input. Lipkin and Hardy (ref. 2) attempted to use computers in the differential diagnosis of hemotological diseases; they coded patients' clinical and laboratory data for application to punch cards. The data contained information from the case history and physical examination, and from peripheral-blood, bone-marrow, and other laboratory examinations.

Logical Concepts in Computer Diagnosis

The next step in prediction of the disease category to which the patient belongs is to attempt to program the computer in a manner that will relate the symptoms to the disease. The physician, in his ordinary attempt to do this, dwells on the patient's symptoms, sex, and age and tries to assign a diagnostic value to each symptom for each disease possibility. He then attempts to relate for each disease the total value of the symptoms for each disease to the totals held in his memory, which were obtained from other patients of the same sex and age range who are known to have the disease. He draws his conclusions and makes his presumptive diagnosis from this information.

In attempts to simulate the physician's thought process by computers, certain common procedures have been followed and certain interesting variations have been attempted. Since the physician asks his memory for reference data, the computer likewise must have a memory. One way to establish such a memory would be, for instance in the case of hemotological-disease prediction, to list the symptoms of the disease from a standard textbook on hemotology. By transfer of the data to cards and their processing in a computer, the symptoms as identified in the patient can be correlated with the diagnostic criteria. Lipkin and Hardy (ref. 2) used this approach and, on the basis of the correlation coefficients obtained, sorted the cases into three groups. Group-I was identical with one disease category; Group-II with several disease categories; and Group-III not with any disease category. For Group-II the indication was that additional information was needed to establish the diagnosis of any one disease. Addition of further information resulted in the symptoms being correlated with only one disease. Further observation of Group-III revealed that more than one hematologic disease was present.

Each item of information previously coded in each of the diseases was assigned a numerical value. On the basis of its contribution to establishment of a diagnosis, the item was assigned a positive, negative, or zero weight. Each item therefore might have a different weight in prediction of each disease. The ratio of the weight of the hospital data to the sum of all weights of the data of the disease was determined. The true disease was considered to be the one that scored highest in terms of weighted averages.

Since this pioneering study in 1957, researchers have become much more sophisticated. Perhaps the most complete and elegant description of what was to become the common approach to the "reasoning process" of the computer, in diagnosis of disease, is by Ledley and Lusted (ref. 3). Under the title "Reasoning Foundations of Medical Diagnosis," they presented the logical, probabilistic, and value-theory concepts, and the conditional probability or learning device that supports the decision of the computer and the interpreter. Their work was aimed at suggestion of a basis from which practical procedures could be (and have been) worked out.

Basic to the reasoning behind medical diagnosis is the concept of probability, since few diagnoses are made with absolute certainty. These logical concepts by Ledley and Lusted start with two sources of information: the symptoms presented by the patient and medical knowledge concerning the symptom relation to certain disease entities. By use of the symbolism common to the language of the propositional calculus of symbolic logic as the basis for communication of the concepts involved in the logical process, patients can easily be classified according to their attributes. Boolean functions are used to classify patients into more than four classes, which would result from dealing with more than two attributes; they are therefore used to express the contents of the logical system which includes symptoms, medical knowledge, and diagnosis. The logical problem is determination of the diseases f. If medical knowledge E is known and if the patient has symptoms G, the prediction is that he has disease f. Ledley and Lusted present this idea in such a way that one need only determine Boolean function f that satisfies this formula:

$$E \rightarrow (G \rightarrow f)$$

They call this the fundamental formula of medical diagnosis.

Application of the logical basis of this diagnostic system requires the display of all combinations of a patient's symptoms. A given patient must fit into one of the mutually exclusive categories of conceivable disease complexes at a time. Each such complex is a possible medical diagnosis. However, in the absence of certainty, one must quantify the probability of the patient having a particular disease with the medium of conditional probabilities. The conditional probability is the probability of a patient having a particular symptom or symptom complex when selected from a population having a particular disease; the mathematics of this approach is a product of Baye's theorem.

Probability enters into the diagnosis as a problem in evaluation of the conditional probabilities for a single patient. Since these probabilities change constantly as further diagnoses are made, the system must be corrected constantly by use of the diagnosed cases as input data for further probability calculations. In this manner the system is constantly improved and kept current with local conditions. Such updating is a relatively simple matter, easily accomplished by a computer-trained person.

In connection with this approach to making computer-aided decisions, value must be considered. Most often mentioned in connection with this problem is the one-person game based on utility theory. Here a mathematical model that fits the operations of the physician, as he makes his decision, is most useful. A good decision is the choice among alternatives that maximizes the probability of achievement of the correct diagnosis. Such strategy is mathematical in form and can be computerized with ease. If one is satisfied with an expected value approach based on a utility theory, the model for decision-making discussed by Von Neumann and Morgenstern (ref. 4) is available.

Crumb and Rupe (ref. 5) suggest a general plan for a logical sequence in development of techniques using computers as diagnostic aids; the following steps are derived from their method:

(1) Test-area selection: One selects a group of diseases that are difficult to diagnose because of similarity of the symptoms.

(2) Test-data compilation: Data relating symptoms to disease are compiled in such a way that statistical calculation of correlation constants is possible.

(3) Development of the correlation technique: Trial solutions are attempted and established for the original data. The technique is developed from them rather than from irregularities that would be occasioned by introduction of new variables. The nonlinearities in preparation and use of the correlation-constant table are tried and accepted as they prove helpful in dealing with further data collected.

(4) Adaptation of data format and computer programs: One selects the appropriate form for input to the computer and writes the program to be used in the computations.

The system yields what are called probability-index numbers. While these indexes do not represent quantitative probabilities, they are used in ranking of the diseases from "most likely" to "least likely" as the true diagnosis for the patient. It has been pointed out (ref. 5) that the big weakness of the technique lies in the fact that the disease list must contain the correct diagnosis.

Crumb and Rupe reported certain computer considerations of importance. The storage requirements of the diagnostic undertaking might be satisfied by a magnetic tape unit, a large-capacity, random-access memory unit such as the IBM Ramac. The total operational time for the computer is approximately 2 min to handle completely a 48×48 matrix. They believed that an inexpensive, medium-size, general-purpose computer with magnetic-tape facility, such as the Bendix G–15, would be satisfactory.

Brodman et al. (ref. 6) did not provide the computer with sets of symptoms for particular diseases; rather they required the machine system to develop empirically its own diagnostic criteria. The computer was provided with a priori information based on the accumulated experience and knowledge of the medical profession and was instructed how to determine diagnostic syndromes from each patient's sex, age, and medical complaint, correlated with the hospital diagnosis; no additional information or assistance was given. This appears to be both a reasonable and an unbiased approach to evaluation of diagnosis by machine.

Because the method of application of conditional probabilities works most easily when one particular disease is dealt with, a rather complicated situation exists when the patient has more than one disease. It is possible to study disease complexes rather than single diseases, nevertheless, by dealing only with the symptoms that have higher incidence within the disease category than in the total population. But, on the whole, few real difficulties have been encountered in application of conditional probabilities in determination of a diagnosis.

Tolles et al. (ref. 7) attempted diagnosis of disease in the cardiovascular system where quantitative data are quite readily available. Their input data to the computer were obtained mainly from the electrocardiogram (ECG, which provided electrical information about the heart), from the ballistocardiogram (for data concerning the mechanical characteristics of the heart), the phonocardiogram (for data concerning the valvular action of the heart), and the arterial pulse wave (for a reflection of the arterial system). The pathological condition chosen for study, in regard to predictability of its diagnosis, was left-ventricular hypertrophy. Complete measurements were made on five heart beats from each of 15 subjects, and the average value of each point was entered in punched cards to serve as a source of input data to the computer. Means, variances, and correlations were computed as statistical measures upon which to base the interpretation.

Computers and Programs in Medical Diagnosis

The above discussion shows two approaches to solution of the relation between symptoms and diagnoses: (1) simple and multiple correlational techniques and (2) conditional probabilities. Either technique requires use of computers for efficient accomplishment. Various programs exist for performance of such correlational procedures on all sizes of computers. A computer as small and as standard as the IBM–1401 or the IBM–1620 is perfectly adequate in performing such computations where the size of the matrix is not more than 30×30. Larger matrices require more passes of the data through the computer and thus more time, so that the efficiency of the process is lessened when a small computer is used.

Brodman (ref. 8) describes the operation of the computer in relating the input information to the diagnosis, and van Woerkom (ref. 9) provides an actual computer program for performance of the Brodman computations. In general the machine has stored in its memory the symptoms as gathered from the CMI and the patient's age and sex. The computer assigns to every complaint a diagnostic value for each disease and corrects this sum for the patient's age. The computer then relates the patient's corrected sum in each disease to the mean of the corrected sums obtained from other patients for whom a valid diagnosis has been obtained. This yields L which is the likelihood of the patient having any given disease. Brodman used the formula

$$L = N_{P50}/N_{m5}$$

where L is the likelihood, N_P is the patient's corrected sum in each disease, N_m is the mean of the corrected sums obtained from other patients known to have the disease, and the lower limit of 5 in the denominator prevents incorrect high values from being assigned to this factor when the mean sum for a disease is low.

Thus the computer output is the likelihood of a patient having each of the diseases (60 diseases in Brodman's study). An L of 35 or more identifies the disease in the patient. An IBM–704 used by Brodman in his computations proved to be quite satisfactory; with it likelihood indexes can be calculated in less than 1 sec.

Tolles et al. (ref. 7, diagnosis of left-ventricular hypertrophy) calculated means, variances, and correlation coefficients; calculations were made from 45 subjects. Because the correlation coefficients were calculated from a 45×45 matrix, the labor would have been enormous without the aid of a computer.

Validation of Computer Diagnosis

The question arises, of course, of just how well the computer diagnosis correlates with a final diagnosis made by well-established methods having the highest validity. All reports in the literature express a great deal of concern in this area and describe in detail the validation procedures and results utilized.

In this Brodman study the diagnosis predicted by computer was compared with the diagnosis made by hospital physicians after eliciting a history and performing physical and laboratory examinations. There was wide variation among diseases in the percentage of cases correctly diagnosed by the computer. Where the CMI elicited many symptoms pertinent to identification of a disease, such as ulcers of the duodenum, as many as 68 percent of the computer's diagnoses were correct. However, where the CMI elicited no pertinent information concerning the symptoms of some diseases, such as benign neoplasm of the skin, no such cases were identified.

The machine and physician experienced in use of the CMI were compared in their ability to diagnose 60 diseases from the CMI data exclusively. The physician correctly identified 43 percent of the cases (other than psychoneurosis) while the machine correctly identified 48 percent, but the difference in percentage was not statistically significant. The physician was clearly superior to the machine in diagnosing psychoneurosis. In some cases the physician could diagnose correctly from the CMI data when the machine could not; however, the machine made few incorrect diagnoses. There was no significant difference between machine and physician in incidence of incorrect diagnoses: 4.9 and 2.0 percent, respectively.

In a cross-validation study in which samples from 1948–49 and 1956 were compared for validity, the results indicated applicability of the technique to a sample different from the one used in its establishment.

When the information input to the computer is systematically biased, the decisions made by the system are incorrect but in a predictable manner. When the input is biased randomly, the computer yields incorrect (but logical) decisions that are unpredictable.

In Rinaldo's study of epigastric pain, the accuracy of the computer in selecting a correct X-ray diagnosis, based on an eight-item questionnaire covering subjective pain symptoms alone, was ascertained by loading of the computer with the subjective symptoms and the X-ray diagnoses of 204 patients. The data were then used by the computer for prediction of the radiological diagnoses of the next 96 patients. The computer correctly identified 73 percent of the patients having hiatus hernia, 69 percent having duodenal ulcers, 27 percent having gastric ulcers, 75 percent having gall stones, 38 percent functional, and 33 percent having gastric carcinomas. Functional disorders caused confusion among the other major diagnostic categories; variability of the data from the patients harmed the validity of the technique. It was suggested that symptom diagnosis could be improved by different weighting of the significance of symptoms and by alteration of one's attitude toward interview data.

The study by Tolles et al. (ref. 7) yielded 69 significant correlations from the 990 calculated; 1 percent (10) of the 990 could be expected to occur by chance alone at the 1-percent level of confidence—the level accepted as significant in this study. Therefore, some of the correlations that could not be explained on a rational basis by the authors should be attributed solely to chance. In order to permit the computer to classify a given individual as normal or as one having a given pathology, Tolles et al. used the means, variance, and correlation coefficients to construct a multidimensional probability model; thus the probability that a given patient would belong to a particular disease or control group could be calculated. The results showed that the ECG, the ballistocardiogram, and the arterial pulse wave could separate the normals from the non-normals, but that the phonocardiogram could not. The ballistocardiogram and the ECG could separate the hypertensive and the valvulitis groups from the others, but the arterial pulse wave could not. The use of such variables improved separation of the normals from the non-normals by 29 orders of magnitude. The authors feel that the method warrants general application and have no doubt of the value of computers for diagnoses.

The validity of the approach used by van Woerkom and Brodman is found in their discussion of their conditional-probability approach. They found that a modified conditional-probability approach, when tested with their sample, produced certain expected inconsistencies; almost always categories were favored having the largest number of attributes with relatively high frequencies. They failed to get probabilities high enough for consideration from categories having small sample sizes and small numbers of attributes of relatively high frequency. They remarked that the latter often would have been identified by a physician, while the machine failed to do so. The validity of their approach rests on the fact that, if a patient's cluster of symptoms resembles the cluster of the average patient, he can be assigned to the correct category and thus validly diagnosed.

The literature makes it clear that the computer's advantages for medical diagnosis are not questioned. Among these advantages are the facts that the complete memory of the computer—but not of the man—is available for making a diagnosis; that the diagnostic time is practically instantaneous once the symptoms input is inserted in the machine; and, through use of an instrument such as the CMI, a presumptive diagnosis may be made even before the physician sees the patient. On the other hand, the most severe limitations of the technique are the scope and quality of the symptom input. If the input is from only the subjective complaints of the patient, the diagnosis varies with the accuracy of that set of data. It is the old story: You get no more out of an analysis than you put into it. A patient has more significant data relevant to his diagnosis than is contained in his medical history. To the extent that these other items, such as laboratory and radiological data, are important to the diagnosis, their exclusion as information renders the diagnosis so much less effective. Presumptive diagnoses by computer, based on both subjective and objective data in sufficient quantities, would of course be completely satisfactory.

Schwichtenberg (ref. 10) has made the point that extension of the use of computers into the life sciences other than medicine, and indeed their use within medicine in such activities as computer diagnosis, depends on development of a standard means of communication; he feels that promulgation of a "current medical terminology" would expedite such communication. The work of the American Medical Association in developing and editing *Current Medical Terminology* for 1963 and 1964, as well as *Current Surgical Terminology* and *Current Medical Surgical Abbreviations*, is either complete or well under way. *Current Medical Terminology* for 1963 (revised in 1964) has already been coded for computers.

MODELS AND SIMULATION OF BIOLOGICAL SYSTEMS

Since the introduction of computers, work on development of mathematical models of biological processes has progressed at a tremendous rate. Complicated systems have been simulated, with the effects of variables within and upon the systems tested, in attempts at more complete understanding of their workings. Although need for such efforts was recognized quite long ago, it was not until the advent of speedy and easy dealing with such complicated models by computers that real progress was made.

Of particular interest in the development of models of biological systems is the analog computer; in general it is smaller and less expensive than its digital counterpart. However, the real reason for use of analog computers in modeling and simulation is that one is essentially able to deal with a direct electrical analog of a physical phenomenon. Taskett (ref. 11) reports that the analog computer permits the research worker to get the feel of a physical system because the results are displayed graphically and faster than is customary with a digital computer.

The analog computer is usually better at performing mathematical operations on measured variables such as are encountered in electroencephalographic and electrocardiographic work.

If the researcher is interested in curve-fitting, density-discriminations, power spectra, and auto or cross correlations, the analog approach is preferable. Thus, for the purposes of model building and simulation of biological systems, the analog computer is preferable. However, when it comes to handling large masses of data from a compilation and statistical-processing point of view, the analog is far inferior to the digital computer. Likewise, when one is not interested in understanding the basic operations of a system but only wants, in a practical sense, to make decisions concerning the operation of a particular subject on a real-time basis and on line, it is better to use an analog-to-digital converter of the biological data and to let the digital computer take over.

We now review current efforts at modeling and simulation, with an eye to understanding of the scope of such modeling efforts, and demonstration of the role played in them by computers.

DeLand (ref. 12) reports an attempt to simulate a large biological system by use of an analog computer—a method based on Gibbs' free-energy hypothesis. A mathematical format was employed, with the actual computations being accomplished by the method of steepest descent. The respiratory function of blood in the human lung was chosen as the subject for the model. The method was based on the postulate that chemical mixtures tend toward a reaction equilibrium that minimizes the potential, or free energy, of the system. The solution of the equilibrium problem consisted of a set of mole numbers that minimized the free-energy function. The time-dependent aspects of the system were also simulated. With use of a digital rather than an analog computer to achieve the same results, it was discovered that the digital approach gave more accurate and reproducible results but had a longer solution time. Since it was desired to simulate the dynamic system in real time, the analog computer was considered better because of its characteristic parallel computation and its fast solution.

An apparatus for simulation of compartmental biological systems was developed by Brownell et al. (ref. 13). The analysis was accomplished by use of an electrical analog. The apparatus was designed for rapid simulation of linear compartmental biological systems and for numerical determination of volume and rate constants from experimental data. The model was applied to two systems: The first dealt with use of the analog as a device for analysis of data, the specific situation chosen being flow and diffusion of the cerebrospinal system; the second dealt with metabolism of iodine by the human thyroid gland. It was found that the voltage curves produced by the compartments bore a scale relation to the activity curves of the biological model.

A good example of how simulation of a system yields new hypotheses, concerning the explanation of phenomena taking place within the system, is reported by Morse (ref. 14). He evaluated the significance of various physiological parameters that contribute to the human ballistocardiogram. Input information to the model consisted of an arterial pulse wave and an arterial ballistogenic function, which is a mathematical representation of the arterial system in which the pulse wave travels. A digital computer performed all the necessary mathematical calculations. Morse found that, although the digital computer is slower than the analog computer in performing such calculations, it has greater flexibility with regard to the input data and the mathematical processes that follow. This investigation showed that the clinical ballistocardiographic abnormalities of *I*- and *J*-wave diminution may be explained as due to changes both in the slope of the rising pulse pressure and in the elasticity of the great vessels. More exact determinations would allow further elaboration of the ballistocardiographic model.

Of particular interest in the area of modeling and simulation is the research of Crosbie et al. (ref. 15). An electrical analog was developed to simulate the physiological responses of man to heat and cold; it was based on the fundamental equation for heat balance, developed to account for heat-loss by radiation, convection, and evaporation. Since physiologic temperature involves at least three types of control mode (proportional, rate, and "on-off"), one must become involved in nonlinear differential equations. The simulator solved such equations in its attempt to predict steady-state situations of rectal temperature, skin temperature, metabolic rate, vasomotor state, and evaporative heat-loss during both rest and exercise. Equations based on the controls mentioned permit simulation of dynamic responses to sudden change in environmental temperature, air velocity, relative humidity, and metabolic rate.

Warner and Cox (ref. 16) investigated the effect on heat rate of stimulation of the sympathetic and vagus efferent nerves to the heart; they hoped for information concerning the nature of the physical and chemical events involved in this transformation and the dynamic and steady-state parameters of this link of the heart-rate control system. Computers were used to increase the accuracy of voltage-to-frequency conversion as well as to perform other important functions in handling of the situation. Differentiations took place on an analog computer. The entire procedure leading to the results was tied to computers at every step, so that the accuracy in simulation and calculation far exceeded what was feasible without computer aid.

Many investigators have made similar use of computerized models. Prominent examples are the work of Clynes (ref. 17) on laws of respiratory sinus arrhythmia, as derived from computer simulation; Pace (ref. 18) on analog-computer simulation of a neural element; and Wood (ref. 19) on mathematical analysis of indicator dilution techniques, where electrical analogs of mathematical models were used to express the distortion of an indicator dilution curve during traverse of a segment of the circulation.

These studies point up the tremendous possibilities of models developed for research purposes. Modeling and simulation are not feasible if the computations are done by hand, nor can one construct a real-time model that does not depend on rapid calculation for its maximum efficiency. Use of the analog computer even in following a simple function is highly desirable. Use of both types of computers jointly for making real-time, on-line analyses of complex functions is indispensable.

Biological systems of interest to medicine are seldom simple. Multivariate situations are routine if a large percentage of all relations within a system are to be understood. It must be remembered, however, that correlation is not proof; while the end product of the *system being simulated* may be the same as that of the *simulation system*, it may have been produced by different means. However, when the researcher treats his results from a simulated system as hypotheses to be verified in the real system that is being simulated, he stands only to gain from the simulation experience.

Great strides have been made in construction of models of biological systems. Equally well known are the methods of setting up of models relating man to his environmental forces. Results yielded concern about (1) the understanding of reactions within a human physiological system, (2) the relation of the action of one physiological system upon another within the human, and (3) the reaction of the total human system to internal and external environmental forces. Besides the nicety of knowing what happens in each of these instances, knowledge can be gained from manipulation of certain variables and combinations of variables to determine their effects on the total-system operation. If the model holds in known instances, and if then its components are experimentally manipulated in previously unknown relations, the effects of these changes in the situational variables upon system performance can be predicted.

CODING OF PHYSIOLOGIC DATA FOR ANALYSIS BY COMPUTERS

Physiological phenomena, picked up directly from subjects by an appropriate lead system, can be recorded on magnetic tape and processed directly by a digital computer. One such system is reported by Taback et al. (ref. 20); they used a corrected orthogonal three-lead system. When a significant cardiac cycle is selected by a technician from an electrocardiographic magnetic tape record, it is automatically sampled at 1-msec intervals, and the numerical values are stored on a magnetic tape in a form acceptable by a digital computer. The researcher is then free to establish various objective analyses for the data and to proceed under each of the programs. The authors offer several reasons for utilizing an automated approach to the study of ECG's. The gist of the reasons follows:

(1) The method provides a completely objective analysis that, of course, does not reflect error due to different interpretations by different readers.

(2) Fidelity of data is maintained in accuracy and in frequency range since no use is made of

direct-writing instruments with their restricted frequency response.

(3) The convenience of magnetic tape for storage and retrieval of large amounts of data, in a form suitable for comparison, is well known.

(4) Because of the high speed of the digital computer, large masses of data can be analyzed statistically with great efficiency; this computer provides an easy and flexible method for investigation of many new and involved approaches to analysis of the data.

Using an IBM–704 a central processing facility accepts the original analog tape recordings and, under the guidance of a medical technician, produces a digital magnetic tape from the information in the form of IBM words. The data are then processed on the 704 according to the program selected by the researcher for the particular analysis.

The speed and efficiency of this objective approach are quite obvious. More research work seems to be needed in determination of criteria for the analyses. Furthermore one should note that, by coupling the results of the analysis with a program for diagnosis of cardiac diseases by computers, a presumptive diagnosis can be achieved. The techniques described here for the cardiac analysis are applicable to other biomedical data.

A great deal of concentrated work on this approach has been supported by the Veterans Administration; it is summarized by Berson et al. (ref. 21). Analog-computer methods for analysis were first considered but then rejected in favor of the digital approach, primarily because of the latter's flexibility. An analog-to-digital conversion of the ECG was made from the original magnetic tape recordings. Specific complexes of the ECG, free of disturbing artifacts, are chosen by an operator, who then presses a button setting the automatic process into motion. The ECG on each channel is converted into digital form and recorded on digital tape at the rate of one conversion per millisecond. The tape is then analyzed by use of the IBM–704. The results of tests of many different programs showed that completely automatic computer analysis of the P–QRS–T complex is highly accurate and of value in large-scale statistical and epidemiological studies. Other analog data, such as phonocardiograms, pulse tracings, and ballistocardiograms, can be analyzed on the same equipment, provided that certain minor modifications are made to the equipment (ref. 22).

Pipberger et al. (ref. 23) have been concerned with the advantages and disadvantages of scalar lead recordings, vector-loop displays, curves of spatial magnitude, orientation and velocity, polar vectors, and eigenvectors. A new method of differential electrocardiography described was based on computer ranges of various diagnostic entities. It was pointed out that the leads that discriminate best between diagnostic groups are obtained by resolution of orthogonal leads. The informational content of three corrected orthogonal leads is comparable to that of the standard 12-lead ECG. Transformation of data appears to be necessary for recovery of the clinical information contained in the 12-lead system.

An automatic method is reported (ref. 24) for processing mass data in clinical medicine, known as FOSDIC (Film Optical Sensing Device for Input to Computers). An analog scanner, activated by a digital computer, automatically reads, recodes, and transcribes the physiological data on magnetic tape in binary language. Research on the system is aimed at its evaluation for reduction of clinical information and correlation of analog records, such as are provided by the ECG, with other clinical data.

Many more biomedical data are produced and normally collected than are required for analysis. As a result some researchers have expressed concern regarding the sampling of data in excess of their needs (ref. 25). This is an important problem that requires solution. Emphasis on this difficulty has been increased by its enhancement of the problem of transmission of medical information over limited-capacity telemetry channels. Bayevskiy (ref. 26) suggests that information theory may solve the problem. Only between 10 and 1 percent of the data collected in some situations may be needed for satisfactory analyses. With an ECG one is dealing with data containing approximately 400 binary units per second. Reduction of this enormous number of bits, by elimination of parts having no practical use, requires special medical devices. By means of such devices the electrocardiographic information can be reduced to a symbolic code.

Transposing Bayevskiy's work on the telemetry system to the problem of coding ECG's for input to a computer, we see that we can apply his principle quite well. He states that the principle of coding an ECG is based on identification of the alpha, beta, gamma, and delta constituents; determination of the integral values of the biocurrents at the output of each filter for a period of 2 sec; and the general integral values of the biopotential in each lead. Thus 20 pulses are available each second, with reduction of the data by 99 percent. Although there is no set of mathematical principles for production of statistical codes for medical data, there is little doubt that the computer itself will be the vehicle for doing the coding as well as the processing of data. Schmitt and Caceres (ref. 27) voice the hope that computers directly accepting biologic data in symbolic or pattern form will eventually evolve.

More research is needed in the area of coding for acquisition, analysis, and presentation of brain-wave data. Computers have been used extensively in development of techniques for dealing with measurements derived from electroencephalographic recordings (refs. 28 and 29, and subsequent work by the same authors).

It is most handy for a physician to have a digital readout of physiological data when his task is to monitor the physiological status of an individual. Siahaya et al. (ref. 30) report a technique that yields digital readout of systolic and diastolic blood pressure, heart rate, and respiratory minute volume, applicable to wireless telemetry from aerospace vehicles; blood-pressure data obtained indirectly in analog fashion are converted to digital information.

Heart-rate QSR complex of the ECG, after passing through wave-shape-recognition and noise-rejection circuitry, is totaled by a digital computer. A voltage-to-frequency converter feeds into a digital counter, which in turn converts analog information to digital and integrates to yield respiratory minute volume. The data are recorded in a predetermined sequence once each minute, and the process is controlled by a programmer.

Hovey et al. (ref. 31) report design of an automatic data-acquisition system that is very useful for minimizing large amounts of biomedical data. The data are digitalized and recorded by a hardware section, reduced by a computer program, and presented in tabular form for analysis. Cady's work (ref. 32) on the development of a computer program for measurement of ECG-wave characteristics is notable in this area.

Glaser's (ref. 33) automatic system for processing microelectrode data is of interest in connection with the uses of computers for analysis of biomedical data. He reports that the data-processor designed and built is useful for transferring the single-unit microelectrode recordings from analog magnetic tape to punched paper tape while preserving the exact time sequences of the neurological events. A Flexowriter can yield a hard copy from the paper tape or produce IBM cards for processing by a computer. Thus one can utilize the digital computer in studying such phenomena as spontaneous activity and adaptive cell processes during and after periods of stimulation.

Such studies as these demonstrate the scope of the digital computer in study of bioelectrical phenomena. In most cases both procedure and apparatus are available, so that maximization of the use of computers in coding and processing of such data is possible.

ANALYSIS AND TRANSMISSION OF PHYSIOLOGICAL DATA

Computers have been used extensively for various mathematical and statistical procedures during the collection and analyses of physiologic data. The neurophysiologist is particularly fortunate in having such a tool; let us cite a specific instance. The computer technique for autocorrelation is of great value; it is a method for comparing a one-time series, such as an ECG, with itself displaced in time. With a lag of 1, 2 . . ., many comparisons are made. If there is a signal in the tracing that is obscured by noise, and if its recurrence is phase locked, the correlation of the tracing with itself is highly positive when the lag introduced equals the repeating period of the signal and whole multiples of the repeating period. The autocorrelation is smaller in value when other lag periods are used. Thus one can pull information from the data that is otherwise obscured. Cross correlations are of equal value. With this method of correlation, two different tracings are compared for determination of whether or not they contain common characteristics.

Vinograd (ref. 34) suggested use of computers for investigation of new physiologic parameters such as rate of change, and rate of rate of change. From this suggestion Townsend (refs. 35, 36) developed the necessary methodology.

In detection of evoked responses, one can detect most efficiently with a computer responses evoked by stimuli (ref. 37). The procedures, including the programs for performance of auto and cross correlations, are now readily available to the researcher and can be routinely employed. Other procedures that are quite helpful and equally available are the methods of phase detection and use of averaging techniques on the computer for analysis of physiological tracings. Adey and Walter (ref. 38) elaborate on these approaches, pointing up the advantages of the computerized techniques for performance of these procedures. In the area of space-flight medical data, Lindsey (ref. 39) has introduced a valuable concept linking statistical procedures with a computerized time-line approach.

Analog-computer techniques have been used in many physiological analyses, as have digital approaches. Typical applications of the analog techniques are those by Murphy and Crane (ref. 40), involving the analog computation of respiratory-response curves; in Randolph's research (ref. 41) into application of analog computers to ESR spectroscopy; and in the studies of Chance et al., with the electric analog computer, of the mechanism of catalase action (ref. 42).

There has been considerable interest in transmission of physiological recordings on an on-line basis. Of course the techniques of telemetry are quite well known and largely standardized. Of more general interest to the medical profession is the effort to use telephone transmission of physiological tracings, such as the ECG, for on-line computer diagnosis. Berson et al. (ref. 43) carefully considered this problem and report success; their transmission system included a regular dial-telephone network for sending and receiving electrocardiographic information. In accuracy the received data were superior to those transmitted by analog systems, since a pulse-code-modulation system, with parity bit transmission and checking, was employed. The data received were immediately analyzed by a digital computer, with verbal delivery of the resultant diagnosis over the same dial-telephone system. Approximately 8 min elapsed between the patient's entry into the ECG room and delivery of a complete P–QRS–T analysis to the transmitting physician.

Another physiological recording was transmitted by a relay satellite. On April 23, 1965, an ECG was transmitted from England via the British transmission system to the relay satellite and thence to the receiving station in New Jersey. It then went by land line to the Mayo Clinic at Minneapolis where it was fed into a computer. From the printout, a diagnosis was made and the results were back in England within 1 min. There appears to be no obstacle to such long-range diagnosis of brain disorders, with proper use of satellite transmission and computer analysis. It is hoped that coupling of computer diagnosis with this system will replace the physician until after a presumptive diagnosis is made.

PHYSIOLOGICAL HUMAN CENTRIFUGE STUDIES

Wood et al. (ref. 44) at the Mayo clinic have been studying the reactions of a physiological system to transient reproducible degrees of stress, as a useful means of discovering the mechanisms of action of the system. Through acceleration they produced reactions in the cardiovascular system that resulted in sudden decrease in arterial pressure at head level, stagnant anoxia of the retina and brain, and hydrostatic effects of acceleration, which alter the ventilation-perfusion ratios in the lungs. A multiplicity of variables had to be dealt with, presenting a task that was ideally suited for computer solution. The hardware used is well described, including the MAYDAC (Mayo analog-to-digital conversion system; ref. 44).

RADIATION TREATMENT

One obstacle to application of automatic calculation of multifield dose distribution is the lack of a workable mathematical description of the percentage-depth dose distribution within a single beam. A formula was found by Sterling et al. (ref. 45) for approximating the percentage-depth distribution resulting from a ^{60}Co beam, of any portal size, at 80 cm SSD. Their equation can be used to yield computer printouts of combined multifield distributions. With a different program it can also print out the combined isodose curves directly to scale, with small error. Thus one can plan and replan treatments according to each individual patient's needs (ref. 46).

MEDICAL LIBRARIES

A good many scientific libraries are computerizing their card catalogs; some produce catalog cards and others have gone to book catalogs. Kilgour et al. (ref. 47) further report that some have gone to an information-retrieval system for catalog-card information that utilizes a large, high-speed computer. Little, however, is being done to computerize the retrieval of catalog and index information except for the MEDLARS system of the National Library of Medicine and the American Society of Metals' information-retrieval system. Both these systems are based on use of magnetic tape. Kilgour et al. also discuss the Columbia-Harvard-Yale Medical Libraries' Computerization Project; its goal is increase in the speed and completeness with which a user obtains catalog and index information in a library. It is hoped that other systems like MEDLARS will be included in the library's facilities. The complete system, including both hardware and software, is hoped to be sufficiently generalizable for use by most conventional libraries. An on-line computer catalog, located at one institution, will allow the processing of requests from various information stations, located at other libraries, by means of a telecommunications system. A random-access memory unit will hold the catalog files. It is planned that the computer will hold the catalog files. It is also planned that the computer programs will be designed for the IBM–1401, 4K-core, two-tape-drive computer; the 1401 is in most general use, being used at five times as many installations as its nearest competitor, the IBM–1620.

There has been some interesting research in the area of "clumping" for associated-document retrieval. A clump, in terms of word association, is a group of objects, from some universe of objects, such that the members of the group (subset) are more closely related to each other than to the rest of the universe to which they belong. The method (ref. 48) involves a statistical technique that can be used for computation of the word associations of use in retrieval of documents from a collection in which the documents are described by index words occurring in the text or in abstracts of the documents. The computerized technique shows evidence of being useful in large-scale associated-document retrieval.

PATIENTS' MEDICAL RECORDS

The need for automatic information-processing of hospital records has been made vividly clear (ref. 49). The population explosion will steadily increase demands on the modern hospital for better facilities, staff, and medical services. The combination of more patients, with better medical coverage, and more medical knowledge will result in much more data that will have to reach quickly the men making the decisions. Furthermore, all medical data on one patient or on all patients—past and present—should be available for different multiple-item correlations to be used for treatment or research. By proper handling of patients' medical records, collected data will be disseminable throughout the hospital as needed; errors will be eliminated, because the information will be transmitted electronically; statistical information regarding the inpatient population will be made available routinely or on demand; the machine will ask for an instruction if one does not come when it is due; it will be possible to select cases as a group for studies, surveys, or research; problems associated with the location, filing, and retrieval of patients' records will be virtually eliminated; and it will be possible to determine trends, and so anticipate conditions developing within a patient or within a hospital.

A list of problems has been published (ref. 49) with which the medical librarian may expect to deal. Among them are definition and determination of satisfactory solutions to medicolegal aspects of automated record systems, establishment of a set of recognition rules to limit access to persons entitled to address the system (for a given purpose or for particular data); and decision on the type of output copy for certain demands.

Of course the first step in conversion from a manual method of handling medical data to an automatic one involving use of computers is performance of a system analysis. Slavin (ref. 50) says that this step entails analysis of the components comprising the existing patient-data system; the clinical folder of the individual patient is the most significant component. Slavin further points out that there are two major types of data contained in the medical record: hard—numerical or identifying data such as for personnel-identification, laboratory-test results, etc.; and soft—evaluative information expressed in narrative form, consisting mainly of physicians' comments. Entering of these data in a computerized system presents a problem, two aspects of which are (1) recording of the essential data and (2) their collection from various locations in the hospital.

The problem of simplification of the collection and recording of data in a standard way has been coped with very well in connection with medical records of astronaut candidates (refs. 51 and 52). Flight surgeons recorded the results of physical examination of the candidates by checking marks on mark-sense cards for reading by a computer; the medical histories and results of laboratory tests could be recorded similarly. Advantages and disadvantages of more than one approach to the card system are discussed (refs. 51 and 52), the conclusion being that despite disadvantages the mark-card system is superior to the usual coded-IBM-card method using work sheets and punch-card operators.

Lindberg (ref. 53) reports on the program surrounding installation of an IBM–1410 computer at a university's medical center. The initial project involved development of a system for reporting all clinical-laboratory determinations through a computer. The data were transmitted to the ward by a preliminary system based on an IBM–1912 card-reader and Teletype Corporation transmitting equipment. Thus the data were already on punched cards and ready for processing by the computer. In a new system under development, the data pass through the computer, before being transmitted, for correction of errors and for general maximization of quality control. Lindberg points out the need for suitable coding procedures for the interrogative medical history and the physical examination. Later he deals further with the technique (ref. 54).

One large-scale effort toward establishment of a computerized system for handling of extensive data and records was made by a large northeastern general hospital. A Hospital Computer Project examined the feasibility of use of a computer to improve patient care and to provide new techniques for research on information in the medical record. The goals of the project were as follows:

(1) Use of a time-shared computer, with remote input-output devices, to increase the rapidity and accuracy of collection, recording, transmission, retrieval, and summary of information

(2) Decrease in the amount of routine paper work required of the nursing staff

(3) Arrangement and consolidation of information for effective utilization by the medical staff

(4) Development of a system that would store large amounts of complex medical information for rapid and easy search and retrieval for facilitation of clinical research

A status report indicated that one of the biggest factors in successful operation of the system is the education of the hospital staff at all levels. A questionnaire, seeking personal reactions of the staff, led to the conclusion that the greater the knowledge and involvement of the participant the more favorable and helpful his attitude. However, there appeared to be a deeper reaction against the total picture of automation. An effective educational program must accompany such an installation if it is to be efficient.

MEDICAL SCHOOLS' COMPUTING FACILITIES

Considerable materials were received from the medical schools contacted during this study; some mailed boxes of publications, descriptions of their facilities, and lists of projects completed by personnel using the facilities. The response in terms of facilities available or planned at the medical schools was so overwhelming that the topic cannot be treated on a school-by-school basis; moreover, some schools placed restrictions on publication of some of their materials. An overall review of the materials received will be presented with the aim of establishing the nature of the average facility now either planned or in operation.

The title of this section, indicating that it deals with computing facilities of medical schools, may give the reader a wrong impression. In some instances the dean of a university's medical school sent materials that reflected the availability of its computing facilities on a university-wide basis, although in fact the university's facilities were available to the medical school.

Generally the medical school lagged behind the rest of the university's schools in the magnitude of its computing facilities. Some medical schools had no computing equipment that they could call their own. The average medical school, if it owned a computing facility independently of the university's other schools and centers, had a small computer such as the IBM–1401 and an odd assortment of analog computers often homemade and hybrid in type. At medical schools at which individuals had special interests in use of computers for modeling or simulation of biomedical process, there was an abnormally high concentration of computer activity, supported by either an exceptional computing facility at the school or a tie-in with a large facility in the vicinity.

Software and Hardware at Medical Schools

Below are listed the computing facilities of one respondent in the study; the school is average in that its hardware is typical of that available in and/or planned for most medical schools:

Available:

1 IBM–1620, 20K
1 IBM–1622 card reader-punch
1 IBM–1443 printer
2 IBM–1311 disk drives
20 IBM–1316 disk packs
Supporting unit-record equipment

On order:

1 IBM–360 model 30E
1 1403 printer
2 1403 disk drives
1 1402 card reader-punch
1 1050 remote terminal

The computer at this medical school is used for business and accounting, research, and teaching; virtually all operations are routine at present. As yet there are no tie-ins with other computing facilities, but a teleprocessing link with a larger computing center is contemplated. There are no special data-handling or data-reduction techniques for physiological data. Analog-to-digital conversion is anticipated in the near future. No computer switching techniques are utilized.

Below are described the computing facilities considered among the best available at larger medical centers where medical research is extensive. The material is taken partly from the publication of a facility that shall not be identified.

The current basic equipment is an IBM–1401 computer with 16 000-character core memory; it includes a 1402 card reader and punch, a 1403 printer, four 729–IV tape units, two 1311 disk drives, and an IBM–1231 Optical Mark Page Reader. Auxiliary card-handling equipment includes a sorter, reproducer, and interpreter and four 026 key punches.

Access is available to a computer center operating a directly coupled system consisting of an IBM–7040 connected to an IBM–7094 computer. The 7040 handles all input/output and buffering, while the more powerful computer, the 7094, compiles, assembles, and executes jobs. The 7094 is a high-speed binary machine with 32 768 addressable locations of 36-bit word size. A large variety of programming languages may be used, and an extensive library of available programs may make it unnecessary for the user to write new programs.

Two source languages are available as part of the operating system of the 1401: FORTRAN and AUTOCODER. Source-language programs in FORTRAN are treated as input data for the FORTRAN compiler program; the compiler is maintained on magnetic tape and is available for use whenever a job is processed. The compiler checks for grammatical errors in the source language and translates the program into object language for the 1401 computer. FORTRAN compilers are available for most computers in use today, so that a program written in this language can be translated and run on another machine with only minor changes.

When a program is to be used repetitively and has been thoroughly checked for ensurance that it operates properly, it can be made available as an object deck, punched in machine code that operates the computer directly. Attempts are made to collect such programs of general usefulness and make them available to others. Some programs are in such general use that they are kept on file on a magnetic-disk unit, and with control cards they can be called for use; the FORTRAN compiler and the AUTOCODER assembler are programs kept available in this fashion.

These general-purpose programs are currently available, all written in FORTRAN:

- 16×16 Correlation
- Direct-difference *t*-test
- Polynomial curve fitting
- *t*-Test/*F*-test
- Chi^2
- Regression lines
- Analysis of variance
- Life table and survival rates
- Mantel/Haentzel chi^2 analysis

This computer center has also obtained the following MEDCOMP programs:

IMP001	Means and standard deviations
IMP004	Linear fit
IMP005	Analysis of variance, one-way
IMP006	Analysis of variance, two-way; no replication
IMP007	Latin square
IMP008	Analysis of variance, two-way; no missing data
IMP009	Analysis of variance, three-way; no missing data
IMP010	Scattergram and grapher
IMP012	Multiple regression
IMP013	Polynomial fit
IMP014	Frequency table generator
IMP015	Marshall test
IMP018	Analysis of covariance
IMP021	Matrix inverter
IMP022A	Expanded histogram
IMP024	Biserial correlation coefficients
IMP025	Generalized accumulator
IMP026	Numerical integrator
IMP027	Two-way analysis of variance, with missing data
IMP028	Analysis of variance, three-way, with missing data
IMP029	Bartlett test for homogeneity of variance
IMP030	Adjustment for heterogeneous variance
IMP031	Correlation coefficients, with missing data
IMP033T	Test-paired and unpaired data
IMP034	Probability chart
IMP035A	Frequency distribution
IMP036	Max-min
IMP037	*Z*-test
IMP039	Exponential fit
ISR002	Log*F*

When one finds specific laboratories as a functional unit within a medical center, he also finds a computing facility that reflects its special needs. If the laboratory is one that makes great use of biomedical data collected in analog form, a computer complex is established to deal with specific as well as general needs. Table 1 is the published list of hardware and software available at or under study for one typical laboratory dealing with biomedical data, and established in connection with a large midwestern medical center.

The general picture of computer use in and around the medical schools of this country is one of great variation. Some schools reported no use of computers for biomedical data. Most schools have either their own computing equipment or ready access to some equipment on the campus; in these cases the amount of research conducted in conjunction with limited and unspecialized computing equipment was often surprising. At some institutions—invariably the larger medical schools near large urban areas and well supported by research grants and other outside sources—one finds highly sophisticated use of computers by workers in the medical area; not in all such cases is there a large computing facility at the school, but usually there is a tie-in with a large nearby computer complex. Such tie-ins bring the best-available computing facilities to the researcher on a time-sharing basis. Smaller schools seem to give more consideration to having a basic computer, such as the IBM-1401 or the IBM-1620, plus a tie-in with a facility having an IBM-7090 or the equivalent; in this way, maximum facilities are available at minimum cost.

TABLE 1.—*Equipment at One Typical Laboratory*

Quantity	Description	Remarks
	Hardware available	
1	IBM-1620 model-II CPU; memory, 20K	1/0 typewriter; automatic floating pt.; auto. divide; indirect addressing
1	Additional 20K core storage	Model 1625
1	Card reader-punch	Model 1622; input speed, 500 cpm; output speed, 250 cpm
1	Magnetic-tape unit with control	Model 7330 and 1921; high-density
3	Card printer-punch	Keypunch machine, model 026
	Hardware under consideration [a]	
4 [b]	Disk storage unit	
6 [b]	Magnetic-tape unit	
	Paper-tape reader-punch	On-line use
	High-speed printer	On-line use
	Digital X-Y plotter	On-line use
100 K [b]	Additional core storage in a system	20K per unit
	Optical input/output subsystem	On-line use
Several	Card printer-punch	Model-026, off-line use
	Accounting machine, IBM-407	Off-line use
	Card sorter	Off-line use
	Paper-tape reader-punch	Flexowriter; off-line use
	Data-medium converter	Off-line use
	Magnetic-tape unit	For analog-data recording
	Analog-to-digital converter	With associated equipment
	Analog computer	
	Software (programming languages) available [c]	
	FORTRAN	Without format
	FORTRAN	With format
	FORTRAN II	
	TABTRAN	
	GOTRAN	
	S.P.S.	
	Machine code	

[a] Not in order of priority.
[b] Maximum.
[c] Users of the facility may prepare their programs in any of these programming languages.

REFERENCES

1. RINALDO, J. A.; SCHEINOK, P.; AND RUPE, C. E.: A Mathematical Analysis of Epigastric Pain. Symptom Diagnosis, vol. 59, no. 2, 1963, pp. 145–154.
2. LIPKIN, M.; AND HARDY, J. D.: Differential Diagnosis of Hematologic Diseases Aided by Mechanical Correlation of Data. Science, vol. 125, March 1957, pp. 551–552.
3. LEDLEY, R. S.; AND LUSTED, L. B.: Reasoning Foundations of Medical Diagnosis. Science, vol. 130, no. 3366, July 1959, pp. 9–21.
4. VON NEUMANN, J.; AND MORGENSTERN, O.: Theory of Games and Economic Behavior. Princeton University Press, 1944.
5. CRUMB, C. B., JR.; AND RUPE, C. E.: The Automatic Digital Computer as an Aid in Medical Diagnosis. Proc. Eastern Joint Computer Conf., vol. 16, 1959, pp. 174–180.
6. BRODMAN, K.; VAN WOERKOM, J.; ERDMANN, A. J., JR.; AND GOLDSTEIN, L. S.: Interpretation of Symptoms with a Data Processing Machine. A. M. A. Arch. Internal Med., vol. 103, May 1959, pp. 776–782.
7. TOLLES, W. E.; STEINBERG, C. A.; CARBERRY, W. J.; AND FREIMAN, A. H.: Experimental Techniques and Results of a Study Using a Computer as a Diagnostic Aid. Trans. N.Y. Acad. Sci., vol. 23, no. 3, Dec. 1961, pp. 246–258.
8. BRODMAN, K.: Diagnostic Decisions by Machine. IRE Trans. Med. Electron., vol. ME–7, July 1960, pp. 216–219.
9. VAN WOERKOM, A. J.: Program for a Diagnostic Model. IRE Trans. Med. Electron., vol. ME–7, 1960.
10. SCHWICHTENBERG, A. H.: What about the Future? Symposium: Computers in Medicine. J. Amer. Med. Assoc., vol. 191, no. 11, 1965, pp. 132–133.
11. TASKETT, M.: Analog Computers in Medical Research and Diagnosis. Arch. Environ. Health, vol. 7, Sept. 1963, pp. 47–56.
12. DELAND, E. C.: Simulation of a Biological System on an Analog Computer. IRE Trans. Electron. Computers, 1962, pp. 17–25.
13. BROWNELL, G. L.; CAVICCHI, R. V.; AND PERRY, K. E.: An Electrical Analog for Analysis of Compartmental Biological Systems. Rev. Sci. Instr. vol. 24, no. 8, Aug. 1953, pp. 704–710.
14. MORSE, R. L.: Significant Physiological Parameters of the Ballistocardiogram as Analyzed by a Mathematical Model. Med. Res. Rept., vol. 1, Jan. 1964, p. 20.
15. CROSBIE, R. J.; HARDY, J. D.; AND FESSENDEN, E.: Electrical Analog Simulation of Temperature Regulation in Man. IRE Trans. Biomed. Electron., Oct. 1961, pp. 245–252.
16. WARNER, H. R.; AND COX, A.: A Mathematical Model of Heart Rate Control by Sympathetic and Vagus Efferent Information. J. Appl. Physiol., vol. 17, 1962, pp. 349–355.
17. CLYNES, M.: Respiratory Sinus Arrhythmia: Laws

Derived from Computer Simulation. J. Appl. Physiol., vol. 15, 1960, pp. 863–874.

18. PACE, W. H., JR.,: An Analog Computer Model for the Study of Water and Electrolyte Flows in the Extracellular and Intracellular Fluids, IRE Trans. Biomed. Electron., Jan. 1961, pp. 29–33.
19. WOOD, E. H.: Abstracts of Current Biomedical Electronic Research Projects. IRE Trans. Biomed. Electron., vol. BME-8, 1961, pp. 265–267.
20. TABACK, L.; MARDEN, E.; MASON, H. L.; AND PIPBERGER, H. V.: Digital Recording of Electrocardiographic Data for Analysis by a Digital Computer. IRE Trans. Med. Electron., vol. ME-6, no. 3, 1959, pp. 167–171.
21. BERSON, A. S.; STALLMANN, F. W.; AND PIPBERGER, H. V.: Digital Computer Analysis of Electrocardiographic Data. ISA Trans., vol. 1, no. 3, July 1962, pp. 287–289.
22. PIPBERGER, H. V.; FREIS, E. D.; TABACK, L.; AND MASON, H. L.: Preparation of Electrocardiographic Data for Analysis by Digital Electronic Computer. Circulation, vol. 21, no. 3, March 1960, pp. 413–418.
23. PIPBERGER, H. V.; BIALEK, S. M.; PERLOFF, J. K.; AND SCHNAPER, H. W.: Correlation of Clinical Information in the Standard 12-Lead ECG and in a Corrected Orthogonal 3-Lead ECG. Amer. Heart J., vol. 61, no. 1, Jan. 1961, pp. 34–43.
24. COSMA, J.; VOLK, M.; GREENOUGH, M. L.; AND PIPBERGER, H. V.: Automatic Method for Processing Mass Data in Clinical Medicine (FOSDIC). Methods Inform. Med., vol. 2, no. 4, Oct. 1963, pp. 125–129.
25. WOODBURY, M. A.: *In* Schmitt, O. H.; and Caceres, C. A.: Electronic and Computer-Assisted Studies of Biomedical Problems. Charles C. Thomas, Springfield, Ill., 1964.
26. BAYEVSKIY, R. M.: Transmission of Medical Information Over Limited-Capacity Telemetry Channels. NASA TT–F–228, 1962, pp. 41–43.
27. SCHMITT, O. H.; AND CACERES, C. A.: Electronic and Computer-Assisted Studies of Biomedical Problems. Charles C. Thomas, Springfield, Ill., 1964.
28. WALTER, D. O.: Spectral Analysis for Electroencephalograms: Mathematical Determinations of Neurophysiological Relationships from Records of Limited Duration. Exp. Neurol., vol. 8, 1963, pp. 155–181.
29. WALTER, D. O.; AND ADEY, W. R.: Spectral Analysis of Electroencephalograms Recorded during Learning in the Cat, before and after Subthalamic Lesions. Exp. Neurol., vol. 7, 1963, pp. 481–501.
30. SIAHAYA, O.; KAHN, A. R.; AND WARE, R. W.: A Digital Readout Technique Applicable to Laboratory and Aerospace Medical Monitoring of Physiologic Data. SAM-TDR-62-139, Feb. 1963, pp. 1–14.
31. HOVEY, W. J.; GILMORE, J. E.; AND KISSEN, A. T.: An Automatic Logging System for Biomedical Test Data. AMRL–TDR–64–50, July 1964, pp. 1–13.
32. CADY, L. D.: A Computer Program for Measurement of ECG Wave Amplitudes, Areas and Durations. 1961.
33. GLASER, E. M.: An Automatic System for Processing Microelectrode Data. IRE Trans. Biomed. Electron., vol. BME-9, no. 3, July 1962, pp. 190–194.
34. VINOGRAD, S. P.: Space Medicine: Scope, Progress and Applicability to Other Fields. Postgrad. Med., vol. 36, no. 2, 1964.
35. TOWNSEND, J. C.: The Measurement of Rate of Change and Rate of Rate of Change of Physiological Data and the Determination of Their Statistical Significance. NASA Memorandum, 1965.
36. TOWNSEND, J. C.: Evaluation of Rate of Change and Rate of Rate of Change as Physiological Parameters. NASA Space Medicine Working Paper, 1965.
37. BRAZIER, M. A. B.: Some Uses of Computers in Experimental Neurology. Exp. Neurol., vol. 2, 1960, pp. 123–143.
38. ADEY, W. R.; AND WALTER, D. O.: Application of Phase Detection and Averaging Techniques in Computer Analysis of EEG Records in the Cat. Exp. Neurol., vol. 7, March 1963, pp. 196–209.
39. LINDSEY, J. F.: Computer Utilization of Time-Line Medical Data from Man in Space Flight. NASA TN D-2695, 1965.
40. MURPHY, T. W.; AND CRANE, T.: Analog Computation of Respiratory Response Curve. Rev. Sci. Instr., vol. 33, May 1962, pp. 533–536.
41. RANDOLPH, M.: Application of Analog Computers to ESR Spectroscopy, Rev. Sci. Instr., vol. 31, Sept. 1960, pp. 949–952.
42. CHANCE, B.; GREENSTEIN, D. S.; HIGGINS, J.; AND YANG, C. C.: The Mechanism of Catalase Action, II. Electric Analog Computer Studies. Arch. Biochem., vol. 37, 1962, pp. 322–339.
43. BERSON, A. S.; STALLMANN, F. W.; BRODERS, J. H.; AND PIPBERGER, H. V.: Telephone Transmission of Electrocardiograms and On-Line Computer Diagnosis. Amer. J. Med. Electron., vol. 4, Jan.-March 1965, pp. 35–37.
44. WOOD, E. H.; SUTTERER, W. F.; MARSHALL, H. W.; AND NOLAN, A. C.: Use of the Human Centrifuge to Study Circulatory Respiratory and Neurologic Physiology in Normal Human Beings and a Description of an Electronic Data Processing System Designed to Facilitate these Studies. AMRL–TDR–63–105, Dec. 1963, pp. 1–24.
45. STERLING, T. D.; PERRY, H.; AND KATZ, L.: Automation of Radiation Treatment Planning. Brit. J. Radiol., vol. 33, 1964, pp. 544–550.
46. STERLING, T. D.; AND PERRY, H.: Planning Radiation Treatment on the Computer. Ann. N.Y. Acad. Sci., vol. 115, July 1964, pp. 976–997.
47. KILGOUR, F. G.; ESTERQUEST, R. T.; AND FLEMING, T. P.: Columbia-Harvard-Yale Computerization Project. Bull. Yale Med. Library, vol. 22, no. 2, Oct. 1963, pp. 7–13.
48. DALE, A. G.; AND DALE, N.: Some Clumping Experiments for Associate Document Retrieval. Amer. Doc., Jan. 1965, pp. 5–9.
49. WILSON, H. H.; AND DEFRANCESCO, H.: Automatic

Information Processing. Med. Record News, April 1963, pp. 60–62.

50. SLAVIN, M.: "EDP" Input Speeds Medical Data Output. Mod. Hosp., 1962, pp. 106–110.

51. SCHWICHTENBERG, A. H.; FLICKINGER, D. A.; AND LOVELACE, W. R., II: Medical Machine Record Cards: Their Development and Use in the Astronaut Selection Program. Bioassay Techniques for Human Centrifuge of Physiological Effects of Acceleration, Lovelace Found. Rept., 1961, pp. 147–174.

52. LOVELACE, W. R., II; SCHWICHTENBERG, A. H.; LUFT, U. C.; AND SECREST, R. R.: Selection and Maintenance Program for Astronauts for the NASA. Aerospace Med., vol. 33, no. 6, 1962.

53. LINDBERG, D. A. B.: A Computer in Medicine. Missouri Med., April 1964, pp. 282–284.

54. LINDBERG, D. A. B.: Electronic Retrieval of Clinical Data. J. Med. Educ., vol. 40, no. 8, 1965.

CHAPTER 2

COMPUTER APPLICATIONS IN THE BEHAVIORAL SCIENCES

S. B. Sells

Man has been both profoundly impressed and baffled by the prodigious advances in science and technology that have at the same time improved and threatened his life since World War II. Whether the path leads to peace and utopian glory or to wanton destruction will undoubtedly depend on how well and how soon man can understand and control himself. Although the sciences of man, and in particular the behavioral sciences, are young and relatively undeveloped, they have started recently to make great forward strides.

This new impetus closely coincided with the availability of high-speed, large-capacity, electronic, digital computers in the 1950's and has kept pace with the phenomenal growth of computer technology in the 1960's. The computer has proved to be a data-processor of vast speed and capacity that has extended research capabilities to a host of significant and previously unassailable problems. In addition, scientists have gained experience and insight into the nature of computers as information-processing systems, and have developed cybernetic models of behavior that promise to illuminate many as-yet-unanswered questions about human behavior.

The purpose of this chapter is to evaluate the impact of computers on the science of psychology, but the treatment of computers as members of a broad family of information-processing systems, observable in nature as well as constructed by man, is emphasized.

NEW RESEARCH CAPABILITY

The new capability that the computer has brought to psychological research, with its data-processing capacity, potential for simulation of behavioral processes and automation of experimental procedures, will change not only the dimensions of research activities but most probably the topology of the entire science. The computer enables the research psychologist to plan for data-analysis, stimulus-generation, real-time control of man-machine systems (including computer-controlled experiments and teaching machines), simulation of self-regulating systems, and learning and problem-solving programs with vastly increased freedom from constraints on implementation, combined with extremely versatile input and output devices. It also gives the research psychologist access to new areas not previously amenable to investigation.

Green (ref. 1) has characterized computers as giant clerks rather than giant brains; in my opinion they can be both. This discussion focuses on the data-processing capabilities of computers programmed by human operators. Whether and under what conditions the brain analogy is justified is considered later. The almost unbelievable computational capability of a modern computer can be appreciated if one example is considered: the time required to (1) read-in magnetic tapes previously encoded with 120 variables per person on a sample of 1000 cases; (2) compute intercorrelations (7140), principal-component-factor analysis, and varimax rotation; and (3) print out results in tabular form. Before electronic computers appeared, this would have been regarded as an "impossible" task; with earlier, smaller computers it may have required several hundred hours, depending on the particular equipment setup. However, with machines such as the CD-1604, the larger CD-3600, or the IBM-7090, this job could be completed in substantially less than 30 minutes.

The enormous, detailed, and tedious labor of arithmetical computation has been reduced by the computer to a rapid routine. At the same time, the numbers of persons sampled, of variables, and of occasions have been vastly extended in relation to processing time. The result is that there is no longer any cogent reason to work with small samples when it is known that the stability of parameters requires considerably larger ones to eliminate relevant sources of variance. Therefore, it is also unnecessary to engage in unreliable research practices on the grounds of expediency. Computers have enabled research workers to approximate infinite series, employ truly multivariate designs, perform matrix algebra and other previously unwieldy computations with ease, and also to approach their scientific problems with previously unknown freedom and power.

Powerful automated routines—In addition to capacity and speed, it is important to realize that the incorporation of simple logical functions permits the programming of automated sequences of operations that overcome some of the most tedious drudgery of extended computations. This can be illustrated by two types of program statements, referred to as the IF and the DO statements, that have become workhorses of computer programmers.

Consider first the following IF statement:

$$\text{IF (Mean } H - \text{Mean } L\text{) 17, 27, 37}$$

This statement, written in FORTRAN, means that the difference, Mean H minus Mean L, is to be computed, and that three branching, alternative instructions are to be followed, depending on whether Mean H is larger than, equal to, or smaller than Mean L. If the difference is negative, the sequence of operations branches to Instruction 17; if it is zero, the operation branches to Instruction 27; and if it is positive, the branching is to Instruction 37. The branching instructions 17, 27, 37 can be located anywhere in the program, and the comparison can be made of any numbers represented by codes in the program, such as Mean H and Mean L in the present example.

An example of a DO statement is next:

$$\text{DO } 25j = 3,\ 50,\ K$$

Let us assume that $K=1$. The computer repeatedly executes all the following instructions up to Instruction 25 in the following way: The first time, operations are performed with j equaling 3; the next time, with j equaling 4 $(3+1)$; the next time, with j equaling 5 $(4+1)$; and so on until j equals 50. After completing the repetitive series of instructions in the DO "loop" with the final computation for $j=50$, the computer resumes the regular sequence of computations following Instruction 25. Each time, through the loop of instructions, computations are repeated with the then-current value of j; the value of K is specified in a separate instruction. Increased flexibility can be gained by placing additional DO loops, and even IF statements, within DO loops.

The IF statement enables the programmer to compare two values (two means, for example) and to proceed in any of three directions depending on whether one is larger than, equal to, or smaller than the other. The DO statement permits automatic repetition of a series of computations before automatic procession to the next phase when that task is completed. Statements such as these, in conjunction with others, permit long and frequently complex sequences of logical decisions and computations. By utilizing numerical coding for alphabetic characters, the computer can handle diverse types of information-processing.

PSYCHOLOGICAL RESEARCH APPLICATIONS

Although psychologists pioneering the new computer applications in psychological research have been relatively few (ref. 2), their accomplishments are impressive both qualitatively and quantitatively. At the same time, thoughtful critics (refs. 3 and 4) have observed that other disciplines have been more vigorous than psychology in employing this powerful tool for study of many aspects of intelligent behavior; they have warned psychologists of the possible consequences to development of their science. Acquisition of computers by universities has been accelerating, however, and it is possible that Baker's (ref. 5) more optimistic view will be realized.

This survey of psychological research employing computers is necessarily brief and is confined to the principal types of application. The discussion is organized under four categories: data processing and statistical analysis, pattern generation and recognition, computer-controlled experiments, and simulation of adaptive behavior. More-detailed accounts are available (refs. 3 to 10).

Data Processing and Statistical Analysis

Computers have increased the magnitude of problems that can be routinely included in multivariate analyses. Thus, whereas a correlation matrix of from 20 to 25 variables was considered large around the end of World War II, the feasible limit for routine operation in the early 1960's was around 200; on the largest computer available, the CD–3600 (maximum capacity, 261 144 48-bit words), it may approach 1000. Along with faster access to storage, faster execution time, and increased accuracy derived from larger word size, the advantages of increased capacity may be expressed by the following rough approximation: an increase in storage capacity from 6000 to 261 000 words (roughly 40 times) increases work capacity by about 1000 times with greater speed and accuracy.

FORTRAN programs have been developed for a wide range of statistical computing, including correlation, multiple regression, canonical analysis, various methods of cluster analysis, factor analysis of R-, P-, and Q-type matrices, and rotation of factor matrices; and also for multidimensional scaling, profile analysis, configural-score analysis, multiple-scalogram analysis, multiple-discriminant analysis, intraclass correlation, various nonparametric analyses, and multivariate analysis of variance (refs. 5 to 8 and 11 to 16). Shepard's development of programs for multidimensional scaling (refs. 17 and 18) has attracted much attention.

Tryon (refs. 15 and 16) has developed an executive program with cluster, factor, and pattern analytic components that can be called in by command at any stage of analysis. His integrated analytic program is capable of extremely versatile multidimensional analysis of correlated data; it was first programmed for the IBM–704 computer and more recently for the IBM–7090. Using such a flexible system the investigator has the option of several alternative procedures at any stage of analysis, and of combining various sequential patterns for empirical comparison of results. Bock (ref. 19) developed a matrix compiler system that was programmed originally for the Univac–1105 and has been reprogrammed for the small LGP–30 at the Psychometric Laboratory of the University of North Carolina. According to Jones (ref. 4), "Each subroutine within the system represents a single matrix operation—addition, subtraction, multiplication, division, inversion, transposition, extraction of diagonal entries, read-in, print-out, etc. Each requires as parameters the order of matrices and the memory location of their initial elements . . . it is arranged [in the Psychometric Laboratory] so that the operator may exercise manual control at the Flexowriter keyboard. The resultant machine resembles a desk calculator for matrix operations." As a computer readily available to graduate students, this appears to be a most valuable training device.

The result of such efforts, which really represent the beginning of revitalization by computer of statistical methodology, has been to make high-powered analysis readily available to almost anyone having funds for computer time. While this result has been a boon to scientific effort, reflected for example by the quality of dissertations using multivariate methods, there have been conspicuous cases of waste and misuse as may be expected. Despite the (relatively small) abuse, however, the computer has greatly advanced knowledge of differences between individuals, particularly in studies of personality dimensions, attitudes, interests, and abilities. Important new studies in personnel psychology, of job components and organization, of personnel selection and utilization, and in simulation of personnel systems also have capitalized on the expanded capabilities provided by computers. Useful summaries of these data-processing uses with extensive references are available (refs. 5 and 12).

Nonquantitative analyses—In addition to statistical applications, computers have facilitated several types of research involving counting and classification of voluminous information, such as in studies of natural-language data. The following examples have been cited (ref. 4) as illustrative of a particular facility of computers:

(1) Word counts, as illustrated by a study (ref. 20) to resolve an authorship dispute

(2) Studies of transitional frequencies between

pairs of word classes in speech from normal and aphasic speakers (ref. 21)

(3) Processing of speech from normals and aphasics to determine functional regularities between frequency of word occurrence and rank popularity of the word (ref. 22)

(4) Grammatical classification of words on the basis of dictionary "look-up" procedures (ref. 23)

(5) Translation from one natural language to another (refs. 24 to 27)

(6) Analysis of syntactic structure (refs. 28 to 30)

(7) Storage and retrieval of texts (refs. 31 to 37)

(8) Content-analysis of text material retrieved (ref. 38)

The study of storage, retrieval, and content-classification of texts is currently of great concern to librarians and bibliographers. In addition to the technical problems per se, there are closely associated and fascinating problems relating to the basis of organization of knowledge.

Here is a good example of a computer's power to virtually plow through a massive accumulation of data to compute summary statistics: This experiment (ref. 39) involved a series of discrimination-reaction-time studies in which six subjects were tested daily, 5 days weekly, over a period. For each subject there were 500 RT's daily; for the entire experiment, more than 15 000 RT's weekly. Each set of RT's was analyzed in 2 to 31 different stimulus categories, and day-to-day variations in central tendency, variance, shape of distribution, error rate, error distribution by stimulus category, and complex special analyses of atypical RT's were computed. The magnitude of the computations was so great that if a computer had not been available the experiments would have been impossible (ref. 39). This is but one example of hundreds that could be cited.

Pattern Generation and Recognition

Stimulus generation—Many kinds of psychological experiments require random selection of stimuli from a known universe. This approach would be ideally applicable to test construction; it has frequent application in perceptual and learning studies. If the universe is large and the lists must be long and numerous, the use of random-number tables for item-selection involves an extremely tedious process. Fortunately random-number generator programs have been developed for computers. If the universe is coded, it is possible automatically to select random lists of any length and to present them in any programmed form of computer output. Such programs can also be modified in various ways; for instance, the program can be random in all respects except that, if a certain stimulus occurs, the probability of a certain other stimulus following it may be made higher than for other stimuli. It is possible to specify the probability of occurrence of any stimulus differentially with respect to other possible stimuli, so that the computer can prepare lists according to desired probabilities. With use of rectangular coordinates, computers can select dots according to two-dimensional locations and thus build two-dimensional configurations of dots and/or lines having systematic, random, or statistically constrained random variations. With the aid of printed, cathode-ray-tube, and auditory outputs much flexibility in stimulus-production can be attained.

Using cathode-ray-tube output to generate bar patterns of dots with different probabilities of occurrence, Green et al. (ref. 40) demonstrated that dot-probability techniques can be used to obscure any type of pattern (e.g., a square) for which a mathematical equation can be written. It is easy to compute whether each dot position in a two-dimensional grid is within the square. Perceptual studies of such stimuli, using photographs of the displays generated on the cathode-ray tube by the computer, have produced intriguing results. "Subjectively it appears that as the probability difference decreases, the presence of a shape is still apparent but the contours cannot be perceived." These techniques can also be used for studying contour-formation as a process.

White (ref. 41) used this technique with a different procedure to study form-recognition. First he generated a clear pattern, 5, using a systematic configuration of dots. The pattern was gradually dissolved in successive manipulations by randomly moving each dot a little each time; this was done by having the computer calculate a small random motion for every dot and then activate a camera to take a picture of the entire display after each set of moves. The result is a series of exposures of the number 5 randomly walking to oblivion. White has used films of this kind, played backward, to determine at what point in the series the figures are recognized.

Pattern recognition—Practical as well as theoretical considerations have motivated intensive study of pattern recognition by machines. Borko (ref. 6, p. 295) has listed several potential applications such as automatic handwriting-recognizers for the Post Office, check-readers for banks and clearing houses, photograph-interpreters for military intelligence and meteorological agencies, and text-readers for the blind; such interests have been responsible for considerable research and development. This problem is fundamental to the study of form-perception as well and has received sophisticated but limited attention from psychologists. Surveys of recent developments in pattern-recognition computer programs and their utility as models for form-perception have been published (refs. 3 and 41). Jones' (ref. 4) significant comments agree with those of Uhr that psychologists may find themselves trailing other disciplines in this area if they do not invest more time and effort on these problems.

According to Uhr (ref. 3), "The problem posed the computer or designer of a computer for 'pattern recognition' (or 'character recognition,' as it is sometimes termed when a specific set of predetermined patterns, usually the alphanumeric symbols as printed in a special type font, is the only set to be processed) is the many-to-one mapping of different inputs into appropriate output sets." Such pattern recognition is performed with disarming ease by human perceivers, as by postmen who daily sort mail accurately that no existing machine could decipher. The output sets referred to by Uhr are the sets into which the human perceiver maps the inputs by grouping things "across an unknown set of geometric transformations and deformations." The pattern-recognition problem involves discovery of the operations that effect this matching, whether by the experimenter in the laboratory or by the computer itself, or by both.

The well-known engineering development for processing of bank checks and identification symbols currently in commercial use are not pattern-recognition devices except in a restricted sense. These techniques involve template-matching (e.g., magnetic-ink grid characters on bank checks) in which prepositioned codes are "read" exactly as punched cards or magnetic tapes. Template-matching devices are limited to the positions and flexibility for which they are programmed, and usually cannot cope with even trivial variations that human perceivers handle intuitively.

Nevertheless, Uhr (ref. 3) reports tremendous and exciting progress toward analytic, sophisticated models that are conceptually related to template-matching. He cites programs that "learn" by storing accumulated "experience" (ref. 42); utilize powerful operators such as "edging" (turning areas into contours), averaging, finding "connectivities" (delimiting figures as opposed to background), and counting (computing area and number of objects) (ref. 43); simulate nets of neuron-like elements (ref. 44); generate their own operators (ref. 45); and function flexibly by means of operators capable of accepting inputs over a range of displacement. Programs such as Rosenblatt's Perceptron (refs. 46 and 47) and Selfridge's Pandemonium (ref. 48) have employed parallel processing rather than serial decisions, thus providing redundancy that facilitates correction of errors.

It is noteworthy that these analytic programs have been found to have neural-net or functional interpretations related to mechanisms of form-perception; according to Uhr they have been instrumental in suggesting physiological and psychological experiments. A most striking instance of this is the fact that Lettvin et al. (ref. 49) made a successful search for straight-line and angle operators, in the visual form-perception of the frog, on the basis of hypotheses generated by Selfridge's Pandemonium model.

It should be apparent that, as pattern-recognition programs advance from the simple template-matching models to the most advanced, complex, and strongest analytic models, the nature of the problem changes correspondingly. The strong analytic models are similar to and in fact are models of cognitive processes. Very frequently they have neurophysiological interpretations, and "A strong argument can even be made for the relevance of pattern recognition for learning and concept formation in machines" (ref. 3).

Computer-Controlled Experiments

With the speed, storage capacity, and input-output versatility that has been described, it has not taken psychologists very long to automate their laboratories. The work of Green and of White (discussed above) on stimulus-pattern generation was a step in this direction. Examples of more-fully automated experimental setups are reported in this section for the purpose of illustrating how far automation can be carried, and also perhaps to project a vision of the psychological laboratory of tomorrow. This section also includes a brief summary of programmed learning developments, which are in fact computer-controlled experiments in their present stage of progress.

Automated experiments—One of the most impressive efforts to use a computer in a highly automated apparatus configuration was a study of auditory-discrimination training (ref. 50). The computer was equipped with a digital-to-analog converter, special output equipment, and two electric typewriters for direct (on-line) input and output. With this setup a complete psychophysical experiment could be run under control of the computer program. The subjects were presented with a sound stimulus that could have any of five values on five dimensions: frequency, amplitude, repetition rate, duty cycle, and duration. The subject's task was to identify the sound and respond by typing a series of numbers on the typewriter. The computer informed the subject, by typing information on the typewriter, whether the response was correct, and, if not, in what respects it deviated from the correct response. Having done this, the computer selected each next stimulus on the basis of the subject's pattern of past responses; which had to be recomputed after each response. In this experiment it was possible to run two subjects at once. The computer kept all records and computed all necessary summary statistics; it also computed auditory wave-form patterns by producing digitized signals corresponding to the desired sounds and then playing these time-quantized data through the digital-to-analog converter.

In this experiment the computer was used to generate stimuli, compute the sequence of their presentation, feed information back to the subject, and analyze the results. It seems only a matter of time before this "unbelievable" setup will be accepted as routine by students in experimental-psychology courses. Yet this is only one example of a class of experiments involving stimulus-generation coupled with random-number generators, digital-to-analog conversion, and special outputs for visual and auditory representation, all integrated with the impressive capabilities of computers for rapid computation and versatile input and output of information.

Jones (ref. 4) has described a series of studies in which the subject's essential task is estimation of an unknown population proportion, given only partial information about its value. The subjects are given a small sample of observations but are also able to profit by prolonged experience with the same parametric distribution of population proportions. The experiment is under computer control. The computer draws each sample of stimuli, prints them out, and requests responses and accepts them on the Flexowriter keyboard; it provides feedback to each subject on performance after each trial and cumulatively. Jones has noted that there is greater stability of performance levels within subjects and less-marked individual differences between subjects in these computer-controlled experiments. This fact may reflect greater control of experimental conditions than in conventional experiments in which a person performs the experimenter's functions. He has also observed the same intense interest on the part of subjects in such experiments as has frequently been reported in programmed learning; apparently this is no longer merely a function of novelty, but rather a matter of greater personal involvement in the task.

A possible disadvantage of the automated laboratory, according to Jones, is the isolation of the investigator, preventing him from making personal observations that might lead to deeper insights concerning his study. However, this need not be true; being free from all the countless demands of operation, the experimenter might be able to use his new freedom more advantageously. It is also important to note that computers are still among the most expensive "gadgets" that have appeared on the budget sheets, and expansion of their use in laboratory work as elsewhere may be slowed by this fact. However, as with all equipment, production costs will undoubtedly fall with improvements in manufacture (as in miniaturization with the availability of transistors) and with mass production.

Automated teaching—An exciting recent development in educational psychology has been the wide interest in learning programs and devices for their operation called teaching machines. Although autoinstruction soared to the status of a fad for a few years and many critics predicted that the "boom" would lead to a premature "bust" (ref. 51), the boom appears to have slowed and a number of substantial research programs have focused serious attention on problems that require investigation.

Many of the advantages of programmed learning are accepted without challenge. Learning programs must be specific and cannot be as loose and rambling as textbooks too often are. Autoinstruction is self-paced and provides individually focused feedback and greater opportunity than group instruction for tailoring of lessons to individual requirements. Questions have been raised about the adequacy of programs, programming principles, and particularly evaluation of progress under autoinstruction as compared with comparable text material. Indeed some of the most vociferous critics of autoinstruction argue that no case has yet been made for the superiority of the program to a good text.

For programs that present their material in a predetermined sequence, it may be difficult to resolve this issue. However, the concept of learning automation has already been extended by the computer to a level of flexibility that clearly goes beyond any simple comparison with good textbooks, and demands comparison only with good teachers. With the computer the sequence of material presented can be adapted to the student's performance to give both immediate and cumulative feedback, to keep records and compute results, and, with special output equipment, to present material in any modality or format desired. Computer control of learning programs permits flexible, automatic operation of a program library and rapid access to a wide range of program material.

Silberman and Coulson (ref. 52) have reviewed the history of computer-controlled teaching machines, which has advanced impressively since the first publication in 1959. The state of the technology is illustrated by the experimental teaching machine built by System Development Corporation (SDC) for a research project (ref. 53); it consists of three components:

(1) Bendix G–15 computer—This is a relatively small general-purpose digital computer with paper-tape input. A bell mounted in the computer frame can be rung under computer control, permitting auditory signals. The computer, as central control unit for the teaching machine, determines at all times during a programmed session the materials to be presented to the student, analyzes students' responses received via the electric typewriter, compares these with stored data, and communicates information to the student.

(2) Random-access slide projector—Developed by SDC staff engineers, it displays instructional materials to the student. It holds up to 600 35-mm slides in 15 magazines of 40 slides each. Selection and projection of slides, each of which holds one item, is under computer control. (Other and similar devices store problem materials in internal computer storage.)

(3) Electric typewriter—It is linked to the computer as on-line equipment and serves as a direct, two-way channel of communication between computer and student.

Learning programs involving communication between experimenter and computer are controlled by paper-tape input after conversion from punched cards to tape. This machine is being used for study of the process of preparation of learning programs: for example, "What decision criteria should be used in determining when to branch a student to less-difficult remediable items?"

Expansion of the computer-controlled teaching machine to accommodate large numbers of students working on a variety of programs appears feasible, and research programs toward this end have been started. The SDC has begun development of an expanded educational facility built around the Philco–2000, a large-capacity digital computer; it is called CLASS (Computer-based Laboratory for Automated School Systems) and will have individual displays and communication input-output devices for each student, as well as monitor displays for teachers. The CLASS facility will also receive, process, and print out data related to registration, attendance, tests, students' educational backgrounds, and other information relevant to a complete educational program. Thus CLASS may serve as a laboratory for the study of not only school learning but also many aspects of an entire school system.

There is little doubt that the hardware for automated school systems is feasible. In fact the development of such facilities is more likely to be slowed by the rate of progress in mastering programming principles for autoinstruction, and producing the programs required for their operation, than by the technical engineering wizardry involved in their design. However, experimental facilities, such as that at SDC and similar ones on several university campuses, must be supported if these goals are to be achieved.

Simulation of Adaptive Behavior

Computers have provided an unexcelled means for building and testing of simulation models of information-processing functions of living organisms on at least two levels: neurophysiological and behavioral. At the first level, programs have been written that simulate or synthesize nerve nets and central information-processing functions of organisms. At the behavioral level, since the early work of Bush and Mosteller (refs. 54 and 55), investigators have focused on problems of molar behavior, such as perception, learning, concept-formation, decision processes, and various complex behaviors including checkers, chess, language-translation, musical composition, and managerial decision making. Models of group and organizational behavior also have been published (ref. 56). Although scientists and scholars have shown concern with these problems for many years, a prodigious volume of work—too extensive and diverse for brief summarization—has appeared within the past 5 years: for example, a survey of computer simulation of cognitive processes (refs. 57 and 58) included over 400 published references, more than half of them written since 1958.

In review of different approaches to these problems one important distinction stands out between those who have interpreted simulation more literally and attempted to match the computer process with the natural system, and those who have concerned themselves only with matching of inputs and outputs, leaving the synthesis of the process to the logic and ingenuity of programmer and computer (refs. 4, 5 and 59 to 61). This distinction is similar in some respects to that between cognitive theory and S–R theory in psychology. Although synthetic programs may be criticized on occasion because they perform human-like functions in nonhuman-like ways, they may also lead to significant insights and hypotheses of heuristic value. Indeed, should such research result only in improving the psychologist's ability to interrogate behavior, it will prove adequately fruitful.

Simulation (and synthesis) research is already beginning to have a healthy impact on psychological thinking and theory. As already frequently noted, computers function in very small, step-by-step processes, and the cold reality of detailed specification, implicit in programming, is lethal to slipshod workmanship and nebulous theory. As an example of this, Baker (ref. 5) has cited Uhr's (ref. 3) comments on the major difficulties encountered in recent efforts to express the theories of even such eminent theorists as Hebb and Hull in testable form. To the extent that the overwhelming volume of theoretical development comes from sources not subjected to such rigorous discipline, this observation may be a forecast of major renovation of psychological theory as computer-oriented research expands.

The interdisciplinary nature of many computer-simulation enterprises is emphasized in White's review (ref. 62) of Rosenblatt's book on perceptrons and simulation of brain mechanisms (ref. 47), which sounds a note of caution regarding the contribution of such synthetic models to psychological understanding. White acknowledged the computer program as a powerful language for the construction of "artificial intelligence" models of many human psychological processes: recognition, problem-solving, rote memory, and the like. However, he warned that "The objective of the whole endeavor is easily lost when the program becomes an end in itself." In White's opinion this appears to some extent to have happened with the perceptron; his evaluation is that it is an "unconvincing neurological model since few of its parameters are firmly rooted in neuroanatomical or neurophysiological data and it offers little to psychologists who expect a model tailored to the known facts of human pattern recognition and discrimination." Although it is probably premature to expect such models to solve all the problems of psychology, it may be recognized that the requirement of interdisciplinary treatment exists, whether accomplished by a more generalized "specialist" or by a team.

Computer Programs for Neural Nets

In 1953 Estes and Burke (ref. 63) published a theory of stimulus variability in learning which has been adapted to a digital-computer program. They hypothesized that on successive learning trials some subset of elements, encoded in the neural net, is consistently reinforced. At the outset of practice all the stimulus elements impinging on the organism are associated with the first successful response. After the second successful response, some of these elements remain associated while others do not. Thus, after a series of trials, a subset of stimulus elements should be sufficiently reinforced to form an S–R connection.

This model is probably adequate for the special conditions of certain controlled learning experiments in the laboratory but is very limited. It is too simple for a perceptual experiment in which successive patterns with no stimulus elements in common may nevertheless properly be placed in the same category. Clark and Farley (refs. 64 and 65) extended the model to a two-stage process in some attempts to simulate neural networks for pattern recognition on a digital computer.

The computer was programmed for simulation of a network of randomly connected nonlinear elements that were assigned parameters to represent thresholds and refractory periods. The network was composed of four groups of elements—two of which represented fixed input; two, output (defined as the difference in the number of elements fired in input and output at any instant). One input consisted of firing of all the elements in the first input group and of none in the second; the other input was the reverse of the first. The system operated according to a rule that an element that received excitation above its threshold would fire and transmit excitation to all other elements with which it was connected. The effectiveness of this excitation (its "weight") was considered a property of each particular connection.

By manipulating these weights after each input presentation, Clark and Farley attempted to determine whether the network could be so organized that a given input could become reliably associated with one of the two outputs. They found that this could be done and that one could classify the inputs in the desired manner significantly beyond chance expectation. They also found that this was possible even when the input patterns contained variable elements (noise). This finding resembles a prediction made in 1949 by Hebb (ref. 66) in discussing the function of cell assemblies in recognition.

More-complex programs involving randomly organized networks, various methods of reinforcing the network, and various methods of organizing and reducing the input signals have been studied. Ashby (ref. 67) and Culbertson (ref. 68) have given highly sophisticated treatment to the problems of simulation of brain and nervous system. However, the foregoing illustrates the logic of an approach to this type of simulation.

Mathematical Models

Several investigators have attempted to formulate principles of neurophysiological functioning and molar behavior in terms of mathematical expressions. This approach predates computers by many years, but has increased in momentum since computers came into use. Uhr (ref. 3) has commented on mathematical analyses of pattern recognition, and his remarks have general applicability.

> These have rarely been programmed and tested, and it seems quite likely that they would not work in the practical situation. They seem of value as they explore mathematics for new approaches, as they classify problems, and as they suggest more elegant and less redundant methods for processing patterns, and especially for making use of the information obtained. But they do not seem to attack the fundamental problem of choosing or discovering the best operators for the appropriate many-one mappings. Rather, they usually address themselves to the question: Given a specified set of operations, what are the best methods for accomplishing them or making use of their results? Or they suggest mathematical methods that have been thoroughly developed, and hence might be powerful tools, if appropriate.

In terms of this distinction, mathematical models belong to the synthetic rather than the simulation approach.

Examples of mathematical models of behavior that have been programmed for computers are many. Uhr has cited approaches, to the pattern-recognition problem, using Fourier analyses (ref. 69), quantum mechanics (refs. 70 and 71), and integral geometry (ref. 72). Problems of optimum coding have been approached in terms of statistics and information theory (refs. 73 to 75), and decision theory (ref. 76). Mattson (ref. 77), extending previous work by Uttley (refs. 78 and 79), has programmed a self-organizing system that will discover the proper Boolean function for linear partitioning of an n-dimensional space into appropriate sets. Sebestyen (ref. 80) developed a method of transforming the space within which input patterns are coded; it was highly successful when applied to speech-recognition.

Behavior-Simulation Models

Behavior-simulation models can be described generally as logical, noncomputational uses of computers. They have been applied at an accelerated rate to problems of perception, learning, concept formation and attainment, memory, tracking behavior, and such associated problems as teaching machines and information-retrieval. According to Baker (ref. 5, p. 568) "The expanse of only a part of the field could be demonstrated by a bibliography of articles on cognitive processes which itself would readily exceed 500 entries."

Most reported psychological simulation studies have used the fifth version of an information-processing language approach, developed by Newell, Shaw, and Simon (refs. 81 to 85), that is widely known by the symbols IPL–V. This is an interpretive programming system that manipulates lists of symbols rather than numbers. An example of this type of program is the Elementary Perceiving And Memorizing (EPAM) program (ref. 86); it is designed to perform rote-memory tasks, such as memorization of lists of nonsense syllables, in a manner simulating the behavior of human subjects. For this task, EPAM depends on a small amount of initial information by which the symbols are paired. This linkage is strengthened during the simulation experiment until the list is "learned" to some criterion. An important comment on this program in comparison with learning by human subjects is that the computer learns any list of symbols, as presented, equally well regardless of their form. Whereas association value of symbols used as stimuli is an important problem in human-learning research, this type of complication is not programmed in EPAM. In more-complex programs, however, computers could be instructed to store past experience, and such complications could be added.

Psychologists have made extensive use of the "thinking aloud" method to develop detailed protocols of subjects' behavior on various problem-solving tasks to be simulated by computers. A recent example (ref. 87) is the use of IPL–V to simulate the problem-solving behavior of a single human subject whose task was to determine the pattern in which four switches were set. The "thinking aloud" protocols were studied along with experimenters' observational notes on the subject's overt behavior; then an IPL–V program was written to simulate the behavior of the subject. Both the human subject and the IPL–V program were then presented with an additional switch-setting problem. The outcome was that both the real subject and the simulated subject generated protocols and problem-solving behaviors that were judged highly similar. This report and a similar one (refs. 59 and 60) describe the experimental procedure and simulation process in painstaking detail. Feldman also considered the criteria of judging of similarity at some length. Minsky (ref. 88) has reviewed the progress with heuristic programs of this type as of 1960.

Edward Johnson,* a student of Jones, is responsible for an important innovation in this area. He analyzed problem-solving performance data for 11 subjects and noted common principles and strategies reflected in their performances. On the basis of these principles he then developed a model such that, when certain individual difference parameters were taken into account, the "style" of problem-solution manifested by the computer program would resemble that of one or another individual subject with a high degree of accuracy.

*Johnson, E. S.: The Simulation of Human Problem Solving from an Empirically Derived Model. Unpublished dissertation, Univ. of North Carolina, 1961.

HOW FAR CAN COMPUTERS SIMULATE BEHAVIOR?

A major concern in the preceding discussion has been elucidation of points of similarity between artificial information-processing systems (computers) and natural systems (behaving organisms). At the same time I have refrained from use of anthropomorphic references, such as brain, memory, thought, and the like, in order to avoid confusion in consideration of the question raised in this final section. It is well known that computer programs, developed or being developed, have performed such "intelligent" acts as proving mathematical theorems, playing games (such as chess) with greater skill than the programmers, recognizing spoken language, translating from one natural language to another, composing music, and even inventing new and better computers.

These accomplishments by computers go beyond mere "giant clerical" computing routines in which the machines do precisely what they are commanded to do. Quite to the contrary they reflect more the quality of giant brains capable of discrimination between alternatives, logical decisions, adaptive change as a result of "experience," and a degree of flexibility that inspires use of the adjective intelligent even in the face of strong inhibition arising from knowledge that the behavior is performed by programmed machines.

The controversy over whether or not machines do indeed "think" is truly a matter of semantics and reflects opinions concerning the distinction between manipulation of symbols and thoughts, as well as philosophical positions vis-à-vis the nature of man. As might be expected, extremely divergent opinions have been expressed by highly qualified behavioral scientists identified with computer research. For example, Borko (ref. 6, p. 20) has stated one extreme categorically: "The computer performs mechanical and electronic operations. It is the human interpreter that thinks"; while in an illuminating address, entitled "How to tell computers from people," Saunders (ref. 89) answered his own question by saying, "You can't."

Between these extremes lies a more pragmatic position that begs the question of a true difference and concentrates instead on the value of simulation as a heuristic approach to the study of behavior. This was stated first, to our knowledge, by Charles Sanders Peirce in 1887 (ref. 90): "Precisely how much of the business of thinking a machine could possibly be made to perform, and what part of it must be left for the living mind, is a question not without conceivable practical importance; the study of it can at any rate not fail to throw needed light on the nature of the reasoning process." Similar opinions have been expressed (refs. 4 and 91 to 93).

For the conservative behavioral scientist, this position may serve as sufficient justification for pursuit of the simulation approach without commitment as to whether or under what conditions a machine may be programmed to behave in a thoroughly human-like way. However, as the effort is extended increasingly to more-and-more-complex behavior, with use of improved equipment of greater storage capacity and flexibility, higher speed, and greater miniaturization, the question of limits keeps reasserting itself. Consideration of this question, moreover, is not merely an amusing exercise, but rather a serious examination of the theoretical end-points of the simulation process.

Ulric Neisser, a leading opponent of the proposition that computers can behave in a thoroughly human-like way, argued that "the root of the difference seems to be more a matter of motivation than of intellect." In support of this assertion he listed these three fundamental and interrelated characteristics of human thought that are conspicuously absent from existing or contemplated computer programs (ref. 94, p. 195):

(1) Human thinking always takes place in, and contributes to, a cumulative process of growth and development.

(2) Human thinking begins in an intimate association with emotions and feelings which is never entirely lost.

(3) Almost all human activity, including thinking, serves not one but a multiplicity of motives at the same time.

In Neisser's opinion these defects are immaterial in "technical applications," such as computation, in which it is of no significance whether the computer (or, for that matter, even the operator) approves or disapproves of the results, so long as they are accurate. However, he is deeply concerned about use of computers in social decisions, "for there our criteria of adequacy are as subtle and as multiple motivated as human thinking itself."

What can be said about these three defects that limit the computer and reduce its "intelligence" ultimately below that of fallible man? These defects are considered briefly in the following sections. The reader is referred to Neisser's competent discussion for defense of his position.

Growth and Development

In the present state of technology, computers have a limited behavior repertoire of instructions to execute and limited capabilities represented by storage capacity and access, speed of execution and transfer of information, input-output versatility, and the like. The computers of the early 1960's represent the second generation of electronic computers, and it is reasonable to expect tremendous extension of these capabilities—roughly comparable to inheritance by subsequent generations of characteristics of a mature living organism. This trend may be regarded as an evolutionary development which, although not "natural," is nevertheless relevant to consideration of development. It is of course true that computer hardware may never go through phases within the operating life of a machine, but one must recognize that a mechanism of "species development" exists.

Next let us consider a form of development that can be related to the operating life of a programmed machine. Such a machine would necessarily be a special-purpose computer (e.g., problem-solver) for the sake of simplicity, but ultimately, if it were endowed with sufficient capability, its functions would be unrestricted. In our present state of ignorance no computer has been given an opportunity to accumulate experience over a long period; the exigencies of laboratory costs and work demands have required erasure of storage after completion of one problem to permit reading-in of the next. However, it is instructive to compare the lifelong exposure to information input and feedback in a human being with the very short-term exposure that computers have thus far experienced. It is contended that self-regulating computer programs, if exposed to "experience" of the order of magnitude of human experience, would be able to show growth in their programmed functions in far less time.

Such growth would of course be represented symbolically in information-processing terms, but, in the simulation frame of reference, this is relevant to the argument. Whereas human growth reflects bodily change and accommodation to the social requirements of a culture in addition to expansion of cognitive capacities, these may be regarded as magnifying the order of complexity of the problem but not changing it quantitatively. This point recalls an observation (ref. 67) that in programming of a computer for simulation of behavior it is as important to specify the environment as the characteristics of the simulated organism. While not feasible in the present state of knowledge, seemingly not beyond the realm of possibility is eventual development of computer programs for simulation of growth experiences of a normal child in his social environment, while at the same time an expanded repertoire of knowledge, skills, attitudes, interests, tastes, and the like is acquired.

Emotions and Feelings

In line with such reasoning it seems equally possible to incorporate in a computer both a storage register for feelings, and output devices for emotional expression. Present concerns with the difficulties of cognitive problems should not obscure the fact that the affective problems are not essentially different but only complicating. Conceptually feelings could be coded in any categories desired, and optimally would be based on current psychological insights. Value systems for filtering input signals could be read into storage so that the computer might decide on the feeling code and classify it in storage accordingly, along with other information. On the output side, emotions could be registered by printing-out of messages, discharge of tubes labeled for various endocrine secretions, rattling of drums, or any means desired.

The important question, of the influence that affective factors would exercise on cognitive functioning, cannot now be programmed for the simple reason that knowledge of this interface of the information-processing functions of the organism is extremely meager. However, this may be an important area for simulation study, for much might be learned thereby. The advantage of study of such problems in information-processing terms is eloquently expressed in the following two statements (ref. 61, p. 362) in relation to a different problem:

> In the discovery of the functional relations necessary for transforming a language input to a language output for any given purpose, it is believed that in symbolic form the functions necessary for the transformation of any other behavioral input to output are being studied also. For example, if in hearing and answering a question, a subject must encode a sequence of sounds, correlate certain aspects of this string with stored information, evaluate the result, and encode it into an appropriate motor output, is it not reasonable to conclude that these procedures are typical of the functions used in reacting to a visually perceived situation?
>
> It soon becomes apparent in the study of language behavior that if a psychologist could understand all aspects of man's use of language, he probably would in the process have developed a complete functional blueprint of the relations holding the vast area between stimulus and response. This functional understanding would be independent of the particular symbolism or coding used by any given sense modality.

Motivation

Certainly any adequate simulation of human behavior must include consideration of motivation. That a motivational system could be represented in information-processing terms seems less important than Hunt's (refs. 95 and 96) extremely significant and exciting prospect that motivational aspects of behavior are implicit in the information-processing activities of the organism. Hunt has analyzed the effects of experience, expressed in terms of a continual interaction process, in determining expectancies, sets, and adaptation levels (to use a few converging terms) and then exploited the activating, directing, and reinforcing effects of incongruity and dissonance as a basis of a motivational mechanism.

Much more must be learned about human motivation before simulation programs will be profitable. Such programs will undoubtedly involve a hierarchically organized set of facilitation and inhibition commands representing interrelated wants, needs, attitudes, interests, values, likes, dislikes, and the like, with variable intensity weights and combinative rules. This problem undoubtedly represents the ultimate complexity discussed thus far.

Symbols versus Ideas

One of the favorite points, made by Borko, Neisser, and others of their persuasion on the man-machine simulation issue, is that the machine can manipulate symbols but has no ideas that are represented by the symbols. In a sense this is true. Nevertheless the very concept of coding of information and of mechanistic processing by manipulation of coded symbols makes this objection irrelevant, for there is little likelihood of any greater degree of ideation in the input, storage, control, processing, or output components of the bioelectric computers than there is in the man-made machines. The nature of conscious awareness and the mechanisms mediating conscious experience are virtually unknown and remain as baffling today as to the first philosophers who attempted to penetrate their mysteries. It is apparent that consciousness occurs, but whether it is an epiphenomenon, like the monitor set in the television studio, or integral to the on-going process is not known. What is known is that associative meaning can be coded and that symbols can represent ideas in information-processing terms without limitation as to intentional or extensional reference if such references are included in the coding instructions.

Total Simulation

This speculative discussion has necessarily taken into account the limitations of present knowledge and equipment. The question has been considered in terms of eventual possibility rather than immediate feasibility. In this frame of reference it is assumed (quite reasonably, in view of the rapid advances of science and technology) that both knowledge and equipment will continue the inexorable march toward realization of the possibilities predicted.

In speculation about future developments it seems appropriate to observe that present discussions of simulation of behavior are segmental, relating to particular behaviors in isolation from the total functioning of the organism. Wooldridge (ref. 97) in his fascinating survey of information-processing functions of living organisms demonstrates that even a simple mammal probably uses hundreds of computers, in a hierarchical organization, for its physiological and behavioral functions. Some of these are analog computers, some are digital, and some systems involve components of both types. These develop in constant interaction with an environment and are modified both by growth and by experience. The organism is constantly programmed by its interactions in the continual environmental encounter.

The hardware of the computer is gross and inefficient in comparison with the delicate microminiaturization and efficient architecture of nature; the physical mechanisms of information-processing are unquestionably different in significant aspects. Yet there seems hardly any limit to the extent to which processes identified in organisms could be simulated in computer programs. As simulation studies expand in complexity and invade broader and more-integrated segments of behavior, they will not only contribute to knowledge but also advance the science of man toward the ultimate goal that Saunders and I believe to be within reach. Of course a simulated man will never be human but perhaps only a remarkable facsimile.

REFERENCES

1. Green, B. F., Jr.: Using Computers to Study Human Perception. Educ. Psychol. Meas., vol. 21, 1961, pp. 227–233.
2. Vandenberg, S. G.; Green, B. F.; and Wrigley, C. F.: A Survey of Computer Usage in Departments of Psychology and Sociology. Behav. Sci., vol. 7, 1962, pp. 108–110.
3. Uhr, L.: "Pattern Recognition" Computers as Models for Form Perception. Psychol. Bull., vol. 60, 1963, pp. 40–73.
4. Jones, L. V.: Beyond Babbage. Psychometrika, vol. 28, 1963, pp. 315–332.
5. Baker, F. B.: Use of Computers in Educational Research. Rev. Educ. Res., vol. 33, 1963, pp. 566–578.
6. Borko, H., ed.: Computer Applications in the Behavioral Sciences. Prentice-Hall, Inc., 1962.
7. Cooley, W. W.; and Lohnes, P. R.: Multivariate Procedures for the Behavioral Sciences. John Wiley and Sons, Inc., 1962.
8. Green, B. F., Jr.: Digital Computers in Research. McGraw-Hill Book Company, Inc., 1963.
9. Suppes, P.; and Atkinson, R. C.: Markov Learning Models for Multiperson Interactions. Stanford Univ. Press, 1960.
10. Ward, J. H., Jr.: Markov Models and Monte Carlo Techniques. Educ. Psychol. Meas., vol. 21, 1961, pp. 219–225.
11. Dixon, W. J.: BIMED: A Series of Statistical Programs for the IBM 709, Part 1. Behav. Sci., vol. 7, 1962, pp. 264–267.
12. Kelly, E. L.; and Lingoes, J. C.: Data Processing in Psychological Research. Computer Applications in the Behavioral Sciences, H. Borko, ed., Prentice-Hall, Inc., 1962, pp. 172–199.
13. Leeson, D. N.; and Dimitry, D. L.: Basic Programming Concepts and the IBM 1620 Computer. Holt, Rinehart and Winston, Inc., 1962.
14. Leone, F. C.; et al.: Abstracts of Statistical Computer Routines. Report no. 1056, Statistical Laboratory, Case Institute of Technology, 1961.
15. Tryon, R. C.: Domain Sampling Formulation of Clusters and Factor Analysis. Psychometrika, vol. 24, 1959, pp. 113–135.
16. Tryon, R. C.: BC TRY System of Multidimensional Analysis. Dept. of Psychology, Univ. of California, Berkeley, Oct. 1962.
17. Shepard, R. N.: The Analysis of Proximities: Multidimensional Scaling with an Unknown Distance Function: I. Psychometrika, vol. 27, 1962, pp. 125–140.
18. Shepard, R. N.: The Analysis of Proximities: Multidimensional Scaling with an Unknown Distance Function: II. Psychometrika, vol. 27, 1962, pp. 219–246.
19. Bock, R. D.: Programming Univariate and Multivariate Analysis of Variance. Technometrics, vol. 5, 1963, pp. 95–117.
20. Mosteller, F.; and Wallace, D. L.: Inference in an Authorship Problem. J. Amer. Statist. Assoc., vol. 58, 1963, pp. 275–309.
21. Fillenbaum, S.; Jones, L. V.; and Wepman, J. M.: Some Linguistic Features of Speech from Aphasic Patients. Language Speech, vol. 4, 1961, pp. 91–108.
22. Osgood, C. E.; and Miron, M. S.; eds.: Approaches to the Study of Aphasia. Univ. of Illinois Press, 1963, pp. 127–130.
23. Jones, L. V.; Goodman, M. F.; and Wepman, J. M.: The Classification of Parts of Speech for the Characterization of Aphasia. Language Speech, vol. 6, 1963, pp. 94–107.
24. Delavenay, E.: An Introduction to Machine Translation. Praeger, New York, 1960.
25. Edmundson, H. P., ed.: Proceedings of the National Symposium on Machine Translation. Prentice-Hall, Inc., 1961.
26. Oettinger, A. G.: Automatic Language Translation. Harvard Univ. Press, 1960.
27. Yngve, V. H.: Computer Programs for Translation. Sci. Amer., vol. 206, 1962, pp. 68–87.
28. Green, B. F., Jr.; Wolf, A. K.; Chomsky, C.; and Laughery, K.: Baseball: An Automatic Question-Answerer. Proc. West. Joint Comput. Conf., vol. 19, 1961, pp. 219–224.

29. HARRIS, Z. S.: Strong Analysis of Sentence Structure. Mouton, The Hague, 1962.
30. SIMMONS, R. F.; KLEIN, S.; AND MCCONLOGUE, K.: Toward the Synthesis of Human Language Behavior. Behav. Sci., vol. 7, 1962, pp. 402–407.
31. BAKER, F. B.: Information Retrieval Based upon Latent Class Analysis. J. Assoc. Comput. Mach., vol. 9, 1962, pp. 512–521.
32. CHEATHAM, T. E., JR.; AND WARSHALL, S.: Translation of Retrieval Requests Counched in a "Semiformal" English-like Language. Commun. Assoc. Comput. Mach., vol. 5, 1962, pp. 34–39.
33. KENT, A.: Information Retrieval and Machine Translation. Vol. III, Pt. 1. Interscience, 1960.
34. KENT, A.: Information Retrieval and Machine Translation. Vol. III, Pt. 2. Interscience, 1961.
35. NEWMAN, S. N.: Information Retrieval, Toward an Ultimate Universal System. Rev. Intern. Doc., vol. 29, 1962, pp. 97–99.
36. SWANSON, D. R.: Interrogating a Computer in Natural Language. Proc. Intern. Fed. Inform. Process. Congr., vol. 62, 1962, pp. 126–127.
37. VICKERY, B. C.: On Retrieval System Theory. Butterworths, London, 1961.
38. STONE, P. J.; BALES, R. F.; NAMENWIRTH, J. Z.; AND OGILVIE, D. M.: The General Inquirer. A Computer System for Content Analysis and Retrieval Based on the Sentence as a Unit of Information. Behav. Sci., vol. 7, 1962, pp. 484–498.
39. SEIBEL, R.: Computer Solutions to Some Noncomputational Psychological Problems. Educ. Psychol. Meas., vol. 21, 1961, pp. 185–201.
40. GREEN, B. F., JR.; WOLF, A. K.; AND WHITE, B. W.: The Detection of Statistically Defined Patterns in a Matrix of Dots. Amer. J. Psychol., vol. 72, 1959, pp. 503–520.
41. WHITE, B. W.: Studies in Perception. *In* Computer Applications in the Behavioral Sciences, H. Borko, ed., Prentice-Hall, Inc., 1962, pp. 281–305.
42. BARAN, P.; AND ESTRIN, G.: An Adaptive Character Reader. IRE Wescon Conv. Record, vol. 4, pt. 4, 1960, pp. 29–36.
43. KIRSCH, R. A.; CAHN, L.; RAY, C.; AND URBAN, G. H.: Experiments in Processing Pictorial Information with a Digital Computer. Proc. East. Joint Comput. Conf., vol. 12, 1957, pp. 221–229.
44. LOEBNER, E. E.: Image Processing and Functional Retina Synthesis. Paper presented at Bionics Symposium, Dayton, Ohio, 1960.
45. UHR, L.; AND VOSSLER, C.: Suggestions for a General Purpose Adaptive Computer Model of Brain Functions. Behav. Sci., vol. 5, 1961, pp. 91–97.
46. ROSENBLATT, F.: The Perceptron: A Theory of Statistical Separability in Cognitive Systems. Report No. VG–1196–G–1, Cornell Aeronautical Laboratory. 1958.
47. ROSENBLATT, F.: Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms. Spartan Books, 1962.
48. SELFRIDGE, O. G.: Pandemonium: A Paradigm for Learning. *In* Mechanization of Thought Processes. Her Majesty's Stationery Office, London, 1959, pp. 511–535.
49. LETTVIN, J. Y.; MATURANA, H. R.; MCCULLOCH, W. S.; AND PITTS, W. H.: What the Frog's Eye Tells the Frog's Brain. Proc. IRE, vol. 47, 1959, pp. 1940–1953.
50. SWETS, J. A.; GREEN, D. M.; AND WINTER, E. F.: Learning to Identify Non-Verbal Sounds, J. Acoust. Soc. Amer., vol. 33, 1961, p. 855.
51. SELLS, S. B.: Personnel Management. Ann. Rev. Psychol., vol. 15, P. R. Farnsworth, O. McNemar, and Q. McNemar, eds., 1964, pp. 399–420.
52. SILBERMAN, H. F.; AND COULSON, J. E.: Automated Teaching. Computer Applications in the Behavioral Sciences, H. Borko, ed., Prentice-Hall, Inc., 1962, pp. 308–335.
53. COULSON, J. E., ED.: Programmed Learning and Computer-Based Instruction. John Wiley and Sons, Inc., 1962.
54. BUSH, R. R.; AND MOSTELLER, F.: Stochastic Models for Learning. John Wiley and Sons, Inc., 1955.
55. BUSH, R. R; AND MOSTELLER, F.: A Comparison of Eight Models. Studies in Mathematical Learning Theory, R. R. Bush and W. K. Estes, eds., Stanford Univ. Press, 1959, pp. 293–307.
56. COLEMAN, J. S.: Analysis of Social Structures and Simulation of Social Processes with Electronic Computers. Educ. Psychol. Meas., vol. 21, 1961, pp. 203–218.
57. SIMMONS, P. L.; AND SIMMONS, R. F.: The Simulation of Cognitive Processes: An Annotated Bibliography. Inst. Radio Engrs. Trans. Electron. Comput., vol. EC–10, 1961, pp. 462–483.
58. SIMMONS, P. L.; AND SIMMONS, R. F.: The Simulation of Cognitive Processes, II: An Annotated Bibliography. Inst. Radio Engrs. Trans. Electron. Comput., vol. EC–11, 1962, pp. 535–552.
59. FELDMAN, J.: Simulation of Behavior in the Binary Choice Experiment. Proc. West. Joint Comput. Conf., vol. 19, 1961, pp. 133–144.
60. FELDMAN, J.: Computer Simulation of Cognitive Processes. Computer Applications in the Behavioral Sciences, H. Borko, ed., Prentice-Hall, Inc., 1962, pp. 337–359.
61. SIMMONS, R. F.: Synthex: Toward Computer Synthesis of Computer Language Behavior. Computer Applications in the Behavioral Sciences, H. Borko, ed., Prentice-Hall, Inc., 1962, pp. 361–393.
62. WHITE, B. W.: Review of Rosenblatt, Frank: Principles of Neurodynamics: Perceptrons and Theory of Brain Mechanisms. Amer. J. Psychol., vol. 76, 1963, pp. 705–707.
63. ESTES, W. K; AND BURKE, C. J.: A Theory of Stimulus Variability in Learning. Psychol. Rev., vol. 60, 1953, pp. 276–286.
64. CLARK, W. A.; AND FARLEY, B. G.: Generalization of Pattern Recognition in a Self-Organizing System. Proc. West. Joint Comput. Conf., March 1955, pp. 86–91.

65. FARLEY, B. G.; AND CLARK, W. A.: Simulation of Self-Organizing Systems by Digital Computer. IRE Trans. Inform. Theory, vol. PGIT-4, 1954, pp. 76–84.
66. HEBB, D. O.: The Organization of Behavior. John Wiley and Sons, Inc., 1949.
67. ASHBY, W. R.: Simulation of a Brain. Computer Applications in the Behavioral Sciences, H. Borko, ed., Prentice-Hall, Inc., 1962, pp. 452–467.
68. CULBERTSON, J. T.: Nerve Net Theory. Computer Applications in the Behavioral Sciences, H. Borko, ed., Prentice-Hall, Inc., 1962, pp. 468–489.
69. GILMORE, H. F.: The Use of the Fourier Transform in the Analysis of Visual Phenomena. Paper presented at Symposium on Pattern Recognition, Ann Arbor, Mich., 1958.
70. GOODALL, M. C.: Performance of a Stochastic Net. Nature, vol. 185, 1960, p. 557.
71. GREENE, P. H.: A Suggested Model for Information Representation in a Computer That Perceives, Learns, and Reasons. Proc. West. Joint Comput. Conf., vol. 17, 1960, pp. 151–164.
72. NOVIKOFF, A.: Integral Geometry, an Approach to the Problem of Abstraction. Paper presented at Bionics Symposium, Dayton, Ohio, 1960.
73. FRANKEL, S.: Information-Theoretic Aspects of Character Reading. Information Processing, UNESCO, 1960, pp. 248–251.
74. PUGACHEV, V. S.: Optimum System Theory Using a General Bayes Criterion. IRE Trans. Inform. Theory, vol. 6, 1960, pp. 4–7.
75. THUN, R. E.: On Dimensional Analysis. IBM J. Res. Develop., vol. 4, 1960, pp. 349–363.
76. CHOW, C. K.: Optimum Character Recognition System Using Decision Function. IRE WESCON Conv. Record, vol. 1, pt. 4, 1957, pp. 121–129.
77. MATTSON, R.: A Self-Organizing Binary System. Proc. East. Joint Comput. Conf., vol. 16, 1959, pp. 212–217.
78. UTTLEY, A. M.: The Design of Conditional Probability Computers. Inform. Control, vol. 2, 1959, pp. 1–2.
79. UTTLEY, A. M.: Temporal and Spatial Patterns in a Conditional Probability Machine. Automata Studies, C. E. Shannon and J. McCarthy, eds., Princeton Univ. Press, 1956, pp. 277–285.
80. SEBESTYEN, G. S.: Recognition of Membership in Classes. IRE Trans. Inform. Theory, vol. 7, 1961, pp. 44–50.
81. NEWELL, A., ED.: Information Processing Language-V Manual. Prentice-Hall, Inc. 1961.
82. NEWELL, A.; SHAW, J. C.; AND SIMON, H. A.: A Variety of Intelligent Learning in a General Problem Solver. Self-Organizing Systems, M. T. Yovits and S. Cameron, eds., Pergamon Press, 1960, pp. 153–189.
83. NEWELL, A.; SHAW, J. C.; AND SIMON, H. A.: Empirical Explorations of the Logic Theory Machine. Proc. West. Joint Comput. Conf., vol. 11, 1957, pp. 218–230.
84. NEWELL, A.; AND SIMON, H. A.: The Logic Theory Machine. IRE Trans. Inform. Theory, vol. IT-2, 1956, pp. 61–70.
85. NEWELL, A.; AND SIMON, H. A.: Computers in Psychology. Handbook of Mathematical Psychology, Vol. 1, R. D. Luce, R. R. Bush, and E. Galanter, eds., John Wiley and Sons, Inc., 1963, pp. 361–428.
86. FEIGENBAUM, E. A.; AND SIMON, H. A.: Performance of a Reading Task by an Elementary Perceiving and Memorizing Program. Behav. Sci., vol. 8, 1963, pp. 72–76.
87. LAUGHERY, K. R.; AND GREGG, L. W.: Simulation of Human Problem-Solving Behavior. Psychometrika, vol. 27, 1962, pp. 265–282.
88. MINSKY, M.: Steps Toward Artificial Intelligence. Proc. Inst. Radio Engrs., vol. 49, pp. 8–30.
89. SAUNDERS, D. R.: How to Tell Computers from People. Educ. Psychol. Meas., vol. 21, 1961, pp. 159–183.
90. PEIRCE, C. S.: Logical Machines. Amer. J. Psychol., vol. 1, 1887, pp. 165–170.
91. VON NEUMANN, J.: The Computer and the Brain. Yale Univ. Press, 1958.
92. TURING, A. M.: Computing Machinery and Intelligence. Mind., vol. 59, 1950, pp. 433–460. (Can a machine think? The World of Mathematics, J. R. Newman, ed., Simon and Schuster, 1956, pp. 2099–2123).
93. BORING, E. G.: Mind and Mechanism. Amer. J. Psychol., vol. 59, 1946, pp. 173–192.
94. NEISSER, U.: The Imitation of Man by Machine. Science, vol. 139, 1963, pp. 193–200.
95. HUNT, J. M.: Motivation Inherent in Information Processing and Action. Motivation and Social Interaction, O. J. Harvey, ed., The Ronald Press Company, 1963, pp. 35–94.
96. HUNT, J. M.: Intelligence and Experience. The Ronald Press Company, 1961.
97. WOOLDRIDGE, D. E.: The Machinery of the Brain. McGraw-Hill Book Company, Inc., 1963.

CHAPTER 3

MEDICAL DATA FROM FLIGHT IN SPACE: OBJECTIVES AND METHODS OF ANALYSIS

Jefferson F. Lindsey, Jr.

The Medical Data Program of the National Aeronautics and Space Administration is designed to meet three main objectives: (1) the safety of the astronauts while in flight; (2) production of new scientific information from the whole space program; and (3) the standardization of all medical data, derived either during flight or on the ground, so that they are in a mutually interchangeable form for computer input and analysis. Success in this program will facilitate international exchange of such medical data and so contribute to the welfare of mankind.

The first objective entails the acquisition and proper utilization of all available medical data bearing directly on the safety of crewmen in flight. Such data must be in a readily interpretable form so that physicians responsible for monitoring of the medical aspects of space missions can use them in assessment of the well-being of the crew and take appropriate action at any moment while a mission is in progress. Thus appropriate data, telemetered during the flight, must be presented to the physician in such a form that he can compare them immediately with medical data previously acquired during both ground-based studies and space missions.

Data pertinent to the second objective also must be in a readily interpretable and standard form for purposes of comparison, interpretation, and prediction. Yet they need not be available for simultaneous readout and immediate application, because they are used for longer-range scientific products applicable to

(1) Advances in medical science and technology
(2) Increase in safety of future crews
(3) More extensive flights
(4) Development and design of spacecraft equipment involving man-machine relations
(5) Improvement of the criteria for selection and training of astronauts.

The third objective—standardization of data for handling by computer—involves the first two objectives since they cannot be accomplished satisfactorily, efficiently, and expeditiously unless the third is met. Accordingly all medical data, both past and future and whether derived during flight or on the ground, must be recorded and prepared on magnetic tape in a standard manner so that they are in a mutually interchangeable form. By use of a proper standardized form or language, these data can be retrieved from computers and brought to bear on specific past problems as well as on current and future problems that may arise during or after future space missions. The computer programs must be prepared in advance for the various graphic, mathematical, and statistical analyses so that interpretation can be immediate.

Consideration must be given to the specific in-flight and ground-based medical data prepared for computer inputs. The in-flight data telemetered to Earth for medical monitoring during the Mercury, Gemini, and Apollo missions are of three types: physiological, spacecraft environmental, and operational performance. The physiological data for each astronaut include electrocardiographic records and data on respiration, pulse, body temperature, blood pressure, etc. The environmental data include measurements of acceleration rate, space-suit inlet and outlet temperatures, carbon dioxide partial pressure, cabin pressure, etc. The operational-performance data have been restricted for the most part to what each astronaut did or said, but more-elaborate monitoring is now being prepared. Aeromedical preparation (ref. 1) for and observations (ref. 2) from the Mercury missions have been reported.

Ground-based medical data are of many types; some of the more important ones relate to mission simulation, the clinical medical history of the astronaut, and the base line. The mission-simulation data for the Mercury, Gemini, and Apollo programs include information obtained during studies in which a space mission was wholly or partially duplicated on Earth; such devices as a centrifuge, a space chamber, a mission simulator, and a procedures trainer were used in addition to immobilization studies. The data acquired during these simulations include those obtained during real space missions. In most cases, however, the former are far more extensive (refs. 3 and 4).

Data in the clinical medical history of each astronaut are from his cumulative record of periodic physical examinations over a number of years (refs. 5 and 6); they may be supplemented by physiological and psychological test data acquired during his selection. Four phases of the astronaut-selection program, as well as a machine-record system to facilitate recording and analysis of medical data, have been described (ref. 5). The third phase of the selection program, including the tests administered and the results obtained, also has been reported (ref. 7). The clinical medical history includes data obtained immediately before and after each space mission; these are necessary for studies of any significant physiological change resulting from the flight (ref. 8).

The base-line data include those available from the literature, indicating norms and tolerances for humans with respect to earlier medical measurements established under specified conditions. Base-line data require little elaboration apart from emphasis on the fact that they are based on a highly selected sample, the astronauts, as well as on general base-line data available in the literature.

A prime objective in establishment of NASA's medical-data program was preparation of all data in a standard, mutually interchangeable form suitable for computer inputs. A question naturally followed: What type of data should be used initially for establishing the pattern to which all other types of data can be related? As a result of a comprehensive study, the in-flight type of medical data was selected for this purpose.

This selection was based on several considerations. First, in-flight data are the most difficult to obtain; once a space mission has started, there can be no turning back, or second attempt, as would be possible in most ground-based physical examinations, tests, simulations, or medical experiments. Second, these in-flight data are highly important and valid for consideration in analyses because they provide information about the precise reactions of crewmen under conditions of real space flight. Third, the in-flight data have a direct bearing on astronauts' safety; since attention must be focused on some selected aspect of such a comprehensive program, the initial emphasis is better placed on data having a direct bearing on safety of flight. Finally the analyses of in-flight data largely serve to indicate future requirements of data from both ground-based studies and missions.

The next questions to be answered were: How can in-flight medical data be prepared for immediate (instantaneous) use by the medical monitors of each mission and for expedient post-flight analyses? And how can all medical data be prepared in a form for computer inputs, incorporating both in-flight and ground-based data?

During the search for answers to these specific questions, the time-line presentation was developed; the rest of this chapter will be devoted to explanation and application of this concept, with attention focused on in-flight medical data. The inductive leap necessary for conception of the way in which this approach can be extended to ground-based medical data will be left mainly to the reader, since such applications are beyond the scope of this chapter. Furthermore I should point out that the time-line-analysis procedures described herein are applicable to many other problems of a situational-analytic type, ranging from problems involving simple human operations in a controlled laboratory environment to those encompassing time-line analyses of aircrew operations while the crew is in the process of flying a supersonic aircraft on a mission lasting many hours (ref. 9). Researchers and scientific investigators are therefore asked to consider the next portion of this chapter with a view toward not only understanding the described methods of analysis of the in-flight medical data but also modification of the methods for application to their particular disciplines.

PREPARATION OF TIME-LINE MEDICAL DATA

General

In the time-line-analysis approach, data sheets are constructed representing successive time periods. The approach can best be presented by describing how the in-flight medical data from the manned Mercury and Gemini flights have been prepared for computer inputs in a standard, magnetically taped format. All relevant information available for a given brief period of time was printed on one data sheet by use of a computer and its associated equipment; this included all available flight information of value to the physician concerning the well-being of the astronaut for the period represented. Since the physician is interested in a composite presentation of all relevant information during any given period, each data sheet included the astronaut's physiological data, spacecraft's environmental data, and astronaut's performance data. Thus the physician can appraise the relations within and interactions among these various factors. Additional data sheets were constructed for consecutive time periods, each of which showed measurements of the same type as those presented on the preceding sheet but of course differing in value because they pertained to different periods.

The requirement for duration of the periods for data sheets was different for various portions of the mission because the physician is interested not only in change per se but also in the rate of change and in the rate of rate of changes of both physiological reactions and environmental conditions. These kinds of changes are generally more rapid during the stressful conditions of the last 2 min before takeoff and during exit and reentry. Generally less stressful portions of a mission occur during weightlessness, the last 1 hour before flight, and after flight. Thus data sheets, each covering the short period of 10 sec, were selected for the stressful portions of the missions; whereas data sheets, each covering a 1-min period, were selected for less-stressful portions. The reason for this selection can be shown more clearly by description of the blocks of data sheets selected for the various types of analyses to be performed (tables 1 and 2).

Description of Blocks of Data Selected

The first block or group of successive data sheets, for each mission, included a series of 15 consecutive 1-min periods covering the time from 1 hour before ($T-60$) to 45 min before takeoff ($T-45$). The first data sheet or tabulation in this group covered the 1-min period starting at $T-60$ and lasting until $T-59$. The next consecutive 1-min period covered the time between $T-59$ and $T-58$; the next, $T-58$ and $T-57$; and so on until 15 consecutive sheets were printed. This block of 15 sheets covered the time between $T-60$ and $T-45$ for each manned space flight (table 1). This procedure can be extended to include additional missions and simulated missions.

The next group or block of data sheets included the series of consecutive 10-sec periods covering the time between $T-120$ sec ($T-2$) and $T-$zero (table 2). This size of sample will be extensible by use of data from future missions, and from past and future simulated missions.

The next blocks of consecutive data sheets covered the time between $T-$zero and the onset of zero gravity (g) in 10-sec periods, since this time was always stressful. The next block of consecutive data sheets covered the time between $T+30$ and $T+45$ in 1-min periods, since these data accompanied weightlessness—less-stressful portions of the missions. To carry this process to its conclusion, a number of selected blocks of consecutive periods were chosen for detailed analyses. These blocks of data covered times when the astronauts were engaged in such functions as exercising (performing identical exercises), resting or sleeping, performing or monitoring retrofire operations, exiting, or landing. The specific periods chosen are shown in detail in table 3.

Study of the information in figure 1 will provide an insight into the manner of comparison and statistical treatment of time-line data. It becomes obvious that data from one mission can be compared with data covering other missions, within certain limitations. Also, measurements taken early during any given mission can be compared with those taken during the latter parts of selected portions of the same mission for assessment of possible changes. Furthermore, when data are prepared in the manner described, the analyses need not be restricted to the period

TABLE 1.—*Matrix of Data Sheets for the 15 1-Min Periods Between T−60 and T−45 Min; × Is One Data Sheet*

Flight	Period														
	$T-60$ to $T-59$	$T-59$ to $T-58$	$T-58$ to $T-57$	$T-57$ to $T-56$	$T-56$ to $T-55$	$T-55$ to $T-54$	$T-54$ to $T-53$	$T-53$ to $T-52$	$T-52$ to $T-51$	$T-51$ to $T-50$	$T-50$ to $T-49$	$T-49$ to $T-48$	$T-48$ to $T-47$	$T-47$ to $T-46$	$T-46$ to $T-45$
1st Suborbital (MR-3)	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
2nd Suborbital (MR-4)	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
1st Orbital (MA-6)	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
2nd Orbital (MA-7)	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
3rd Orbital (MA-8)	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
4th Orbital (MA-9)	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
1st Gemini (GT-3)	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
2nd Gemini	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
Other	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×

TABLE 2.—*Matrix of Data Sheets for the 12 10-Sec Periods Between T−120 Sec and T−Zero; × Is One Data Sheet*

Flight	Period											
	$T-120$ to $T-110$	$T-110$ to $T-100$	$T-100$ to $T-90$	$T-90$ to $T-80$	$T-80$ to $T-70$	$T-70$ to $T-60$	$T-60$ to $T-50$	$T-50$ to $T-40$	$T-40$ to $T-30$	$T-30$ to $T-20$	$T-20$ to $T-10$	$T-10$ to $T-0$
1st Suborbital (MR-3)	×	×	×	×	×	×	×	×	×	×	×	×
2nd Suborbital (MR-4)	×	×	×	×	×	×	×	×	×	×	×	×
1st Orbital (MA-6)	×	×	×	×	×	×	×	×	×	×	×	×
2nd Orbital (MA-7)	×	×	×	×	×	×	×	×	×	×	×	×
3rd Orbital (MA-8)	×	×	×	×	×	×	×	×	×	×	×	×
4th Orbital (MA-9)	×	×	×	×	×	×	×	×	×	×	×	×
1st Gemini (GT-3)	×	×	×	×	×	×	×	×	×	×	×	×
2nd Gemini	×	×	×	×	×	×	×	×	×	×	×	×
Other	×	×	×	×	×	×	×	×	×	×	×	×

TABLE 3.—*Types of Activities*

Criticalness (C)	Difficulty (D)	Duty (T)	Procedure (P)
1C: Highly critical	1D: Very difficult	1T: New (performed in spacecraft only)	1P: Active
2C: Medium-critical	2D: Medium-difficult	2T: Revised (combination of old and new)	2P: Passive
3C: Noncritical	3D: Easy	3T: Old (previously performed in aircraft)	3P: Concurrent
			4P: Shared (with ground personnel)

	15 MINUTE PERIOD :: 1 MIN SAMPLES AT A COMMON TIME BETWEEN T-60 MIN AND T-2 MIN	T-2 MIN TO LIFTOFF :: 10 SEC SAMPLES	LIFTOFF TO ZERO g :: 10 SEC SAMPLES	ZERO g TO +15 MIN :: 1 MIN SAMPLES	ZERO g +30 MIN TO ZERO g +45 MIN	PERIOD OF ASTRONAUT ACTIVITY	SLEEP	RETRO SEQUENCE -90 MIN TO -75 MIN BASED ON MA-6 & MA-7	RETRO SEQUENCE -90 MIN TO -75 MIN BASED ON (MA-8)	RETRO SEQUENCE -90 MIN TO -75 MIN BASED ON (MA-9)	RE-ENTRY SEQUENCE -15 MIN TO .05g BASED ON MA-6 & MA-7	RE-ENTRY SEQUENCE -15 MIN TO .05g BASED ON MA-8	RE-ENTRY SEQUENCE -15 MIN TO .05g BASED ON (MA-9)	.05g TO LANDING	POST-LANDING
MR-3	15 MIN DATA CONTINUOUS 15 TABS	2 MIN DATA 12 TABS	2 MIN 40 SEC DATA 16 TABS	ONLY 5 MIN OF DATA 5 TABS										00 07 50 TO 00 15 37 47 TABS	LANDING+5 MIN 1 MIN SAMPLE 5 TABS
MR-4	"	"	"	ONLY 5 MIN OF DATA 5 TABS										00 07 50 TO 00 15 37 47 TABS	"
MA-6	"	"	5 MIN DATA 30 TABS	15 MIN DATA 15 TABS	15 MIN DATA 15 TABS	00 25 00 TO 00 30 00 5 TABS		3:09:08 TO 3:24:08 15 TABS			4 29 03 TO 4 44 03 15 TABS			4 44 03 TO 4 55 23 68 TABS	"
MA-7	"	"	"	"	"	3 52 00 TO 4 07 00 15 TABS		"			"			4 44 03 TO 4 55 57 72 TABS	"
MA-8	"	"	"	"	"			"	7:22:00 TO 7:37:00 15 TABS		"	8 46 40 TO 9 01 40 15 TABS		9 01 40 TO 9 13 11 70 TABS	"
MA-9	"	"	"	"	"	2 19 00 TO 2 34 00 15 TABS	13 59 00 TO 15 03 00 27 TABS	"	"	32 29 30 TO 32 44 30 15 TABS	"	"	33 53 36 TO 34 08 36 15 TABS	34 08 36 TO 34 19 49 68 TABS	"
MA-9									25:00:00 TO 25:15:00 15 TABS						"
TOTAL	90 TABS	72 TABS	152 TABS	70 TABS	60 TABS	35 TABS	27 TABS	60 TABS	45 TABS	15 TABS	60 TABS	30 TABS	15 TABS	372 TABS	30 TABS

TOTAL, 1133 TABS

FIGURE 1.—Example of Mercury biomedical-data requirements; tabs, tabulations (data sheets).

represented by each data sheet—10-sec or 1-min periods. If, for example, one has 15 consecutive 1-min data sheets, the computer can be programmed to treat these data as either a consolidated or a segmented block of data of any desired length in minutes or fractions of minutes, such as 15, 10, 5, 3, 1, $\frac{1}{2}$, $\frac{1}{4}$, or $\frac{1}{8}$; the same is true of the 10-sec-period data sheets.

Description of Data-Sheet Content

Detailed examination of the specific information included on each data sheet is now in order; this will be accomplished in conjunction with illustrations of two types of data sheets for selected 10-sec periods (figs. 2 and 3). Data sheets for the 1-min periods were identical in format except that no acceleration data are included, since the 1-min periods are applicable to times of the mission preceding flight, during weightlessness, or after flight, when no acceleration forces were present. Therefore, for all practical purposes, the following discussion of the

TIME-LINE DATA

PROJECT MERCURY
MERCURY ATLAS NO.

MISSION ELAPSED TIME = 00,00,20 TO 00,00,30
DATE

HEART RATE BEATS/MIN	STANDARD SCORE	RESPIRATION RATE BREATHS/MIN	STANDARD SCORE	ACCEL. Z-AXIS G-S	STANDARD SCORE	SUIT-IN TEMP DEG F	SUIT-OUT TEMP DEG F	CO2 PARTIAL PRESSURE PSIA	CABIN PRESSURE PSIA
115.60	38.20	16.90	36.40	1.77	50.30	65.22	84.52	.0000	15.245
112.14	36.79	15.78	35.39	1.90	50.96	65.18	84.46	.0000	15.213
112.14	36.79	22.72	41.63	1.77	50.30	65.22	84.52	.0000	15.198
113.85	37.49			1.71	50.00	65.26	84.58	.0000	15.229
115.58	38.19			1.71	50.00			.0000	15.229
115.60	38.20			1.64	49.65			.0000	15.245
113.85	37.49			1.37	48.28			.0000	15.276
115.60	38.20			1.90	50.96			.0000	15.198
112.14	36.79			2.43	53.65			.0000	15.198
113.85	37.49			2.04	51.67			.0000	15.198
113.85	37.49			1.71	50.00			.0000	15.174
115.58	38.19			1.90	50.96			.0000	15.158
119.28	39.69								
121.18	40.46								
121.18	40.46								
123.20	41.28								
121.18	40.46								
119.26	39.68								
123.17	41.27								
123.20	41.28								
						**** MEAN		**** MEAN	
						65.22	84.52	.0000	15.216

******	******	******	******	******	******
MEAN =	117.07	MEAN =	18.46	MEAN =	1.82
SIGMA=	3.96	SIGMA=	3.72	SIGMA=	.25
VARIANCE=	15.68	VARIANCE=	13.88	VARIANCE=	.06

ACTIVITY

1. PLANNED START BACKUP CLOCK.
2. INDICATED START BACKUP CLOCK.

COMMUNICATIONS

00,00,20 CC MARK.
00,00,23 P ROGER. AND THE BACKUP CLOCK IS RUNNING
00,00,25 CC ROGER. YOU LOOK GOOD HERE, GORDO.
00,00,27 P ROGER. FEELS GOOD, BUDDY.
00,00,29 CC GOOD SPORT.

FIGURE 2.—Exemplary data sheet for 10-sec period.

10-sec type of data sheet will suffice; the discussion is keyed to figures 2 and 3.

Heading—The heading of each data sheet contained information identifying the data as time-line data and indicating the mission from which they were taken, the mission's elapsed time, and its date. The time-line data (fig. 2) were taken from a Mercury-Atlas mission, MA–9; the mission's elapsed time was the time between 20 and 30 sec after takeoff on May 15, 1963. The reason for including the information presented is obvious: each data sheet must be an identifiable entity in itself to be used by the computer for analytic purposes.

Heart rate—The rate of the astronaut's heart beat in beats per minute (R to R) is shown for each beat that occurred during the 10-sec period represented. In the example there were approximately two beats per second since the mean for the 10-sec period is very nearly 120; more precisely it was 117.07. Since the heart does not beat at a constant rate, it was possible to solve for the standard deviation (sigma) and the variance for the period represented. The second column shows the standard score for each beat represented in the first column. These standard (z) scores were calculated by using all the heart-rate data points within a predefined period for any given mission (e.g., all data points during the combined periods of exit and reentry), and then finding the mean and standard deviation of the distribution of these data. Thence the z-score was calculated for each heart-rate data point in column 1. Since z-scores contain negative numbers, a new distribution (standardized score) was formed with a mean of 50 and a standard deviation of 10 by multiplying each z-score by 10 and adding 50. The standardized scores for any given mission in progress cannot be computed instantaneously as the mission progresses since all data points in the distribution (exit and reentry, in this example) must be known before the calculations can be made. However, the standardized scores for each heart-rate data point can be calculated instantaneously if the distribution has already been established during previous missions. Details of methods for these calculations are available (refs. 10 to 12).

Again the reasons for obtaining and printing digital heart-rate data on each data sheet are obvious, but the calculation of means, standard deviations, and standard scores requires at least brief justification. The means and the standard deviations can be graphically printed by the computer

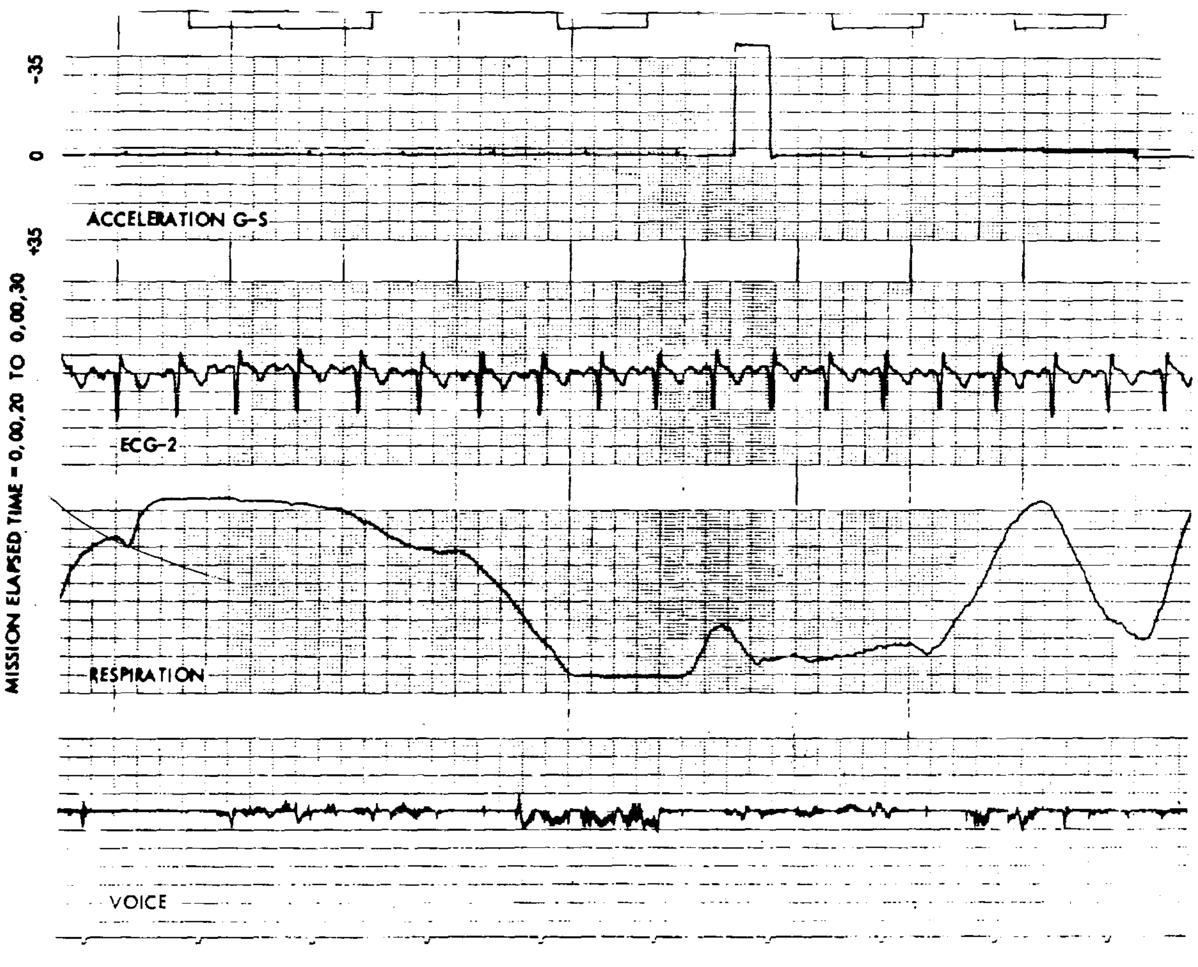


FIGURE 3.—Exemplary data sheet for 10-sec period; ECG, electrocardiogram.

for presentation to the physician as the mission progresses, to indicate trends or patterns that may be developing in heart rate. These statistics can also be used in connection with various graphic, mathematical, and statistical analyses as will be explained later. The standard scores, based on a mean of 50 and a standard deviation of 10, provide the physician with definite information as to how near normal is the heart's beat in comparison to what should be expected at any given time and in any given circumstance for the astronaut(s) participating in the mission. For example, if the standard score is 50 (the mean), the physician knows that this is the average rate for the astronaut when all heart beats to be expected during exit and reentry are considered. If the standard score is 40, he knows that this particular beat is 1 standard deviation below the mean. Additionally by converting to standardized scores instead of using z-scores, the need to work with negative numbers is eliminated; this is generally convenient and sometimes necessary for certain types of analyses. Finally conversion of not only heart-rate but also respiration-rate and acceleration-rate data to standard scores makes possible comparison of these diverse types of measurements by use of a standard base (standard score); accordingly various other analytic techniques become available for use.

Respiration—The respiration rate, in breaths per minute, for the 10-sec period is shown (fig. 2) along with the standard score for each entry. The standard scores are based on all respiration-rate-data points covering the same period of the mission that was used to calculate the heart-rate standard scores—in this instance, exit and reentry. Also, as in the case of heart-rate data, the mean and standard deviation are shown for the respiration data for the 10-sec period represented. The rationale for inclusion of the foregoing respiratory information in each data sheet was similar to that previously described; when using several consecutive data sheets, one can begin to identify existing trends, patterns, and relations.

Other measurements—Other physiological measurements mostly of a periodic nature were taken but are not shown since they are not applicable to the 10-sec-period data sheet used here for illustrative purposes.

Environmental measurements—These measurements include acceleration, space-suit inlet and outlet temperatures, carbon dioxide partial pressure, and cabin pressure (fig. 2). The data sheets included rate information for each of these parameters, plus the mean, standard deviation, and standard scores for acceleration data. Only the mean was calculated for the other environmental measurements. It can be seen now that trend information for environmental measurement becomes available, as do data pertaining to relations both within and among physiological and environmental aspects of space flight.

Activity—One type of operational-performance data of a continuous nature, readily available for incorporation in the selected data sheets, was a description of both planned and actual activity of the astronaut during the period represented. The astronaut was supposed to and did start the backup clock during the 10-sec period (fig. 2).

By recording the planned as well as the actual activity one can tell whether the astronaut is on, ahead of, or behind schedule and assess the implications thereof. However, in order to utilize these data digitally for analytic purposes in relation to physiological and environmental data, it was necessary to construct operational definitions for each type of activity. These definitions were subdivided into four areas having assigned numerical values or weights (table 3).

Voice—Another type of indirect indicator of operational performance, as well as of the well-being of the astronaut, was the voice data; they were supplemented by communications with ground crewmen. Figure 2 shows the communications between ground crewmen and the astronaut during the 10-sec period.

Many types of analysis can handle voice data, ranging from assessment of the state of the astronaut's alertness (and probable relations associated therewith) to analyses pertaining to speech processes, audiology, and information-processing (ref. 13). The physician can gain considerable information about the condition of the astronaut merely by listening to his voice and conversing with him. Here again though, for use of voice data in the computer, voice content must be digitized. Therefore, operational definitions are required encompassing areas such as probable attention, joy, fatigue, confusion, and relief. (These terms originated with L. V. Surgent; see Chapter 10 for details.) Factors considered in classification according to these areas include voice pitch, timbre, and speed, and in some cases the quickness of the astronaut's response to questions or instructions received from the ground. These questions and instructions, however, must be weighed with extreme care because a response may not be immediate because of overriding operational functions in which the astronaut may happen to be engaged at the time. Consideration must also be given to the fact that astronauts are a highly selected group with exceptional ability; they are conditioned to excitement and stress. Therefore, very little degradation in their voice content may be more significant than greater change for the normal population.

Wave-train data—These consisted of analog readouts of the electrocardiogram (ECG), respiration, acceleration, and voice (fig. 3). These are required for two main purposes: to check the correctness of the digital data on the data sheet and to use in connection with certain types of pattern and wave-form analyses.

TYPES OF ANALYSES

General

When data thus prepared in time-line format are available on magnetic tape, many types of analysis can be performed that apply directly to the safety of astronauts in flight and to scientific products derivable from such medical data. Several aspects relative to these analyses, together with limitations, will be discussed now.

Graphic Analyses

One-dimensional graphs of the means, variances, and standard scores can be plotted for comparable time periods for each astronaut on all measurements or for all astronauts on one selected measurement. For example, the graph of heart-rate for the 5 min immediately after takeoff consists of 30 data points—one for each 10-sec period shown on each data sheet for 5 consecutive minutes. Two-dimensional graphs also have been constructed showing, for example, heart-rate on the ordinate and acceleration on the abscissa. Additionally three-dimensional graphs can be constructed with various colors used for the third dimension.

Although most of these graphs can be plotted by the computer in real time as a mission progresses, this information must be compared with earlier data so that it may be meaningful to the medical monitors. Therefore overlays need to be used so that graphic information concerning the current mission can be superimposed on previously constructed graphs based on time-line data acquired from either completed or simulated missions. Overlays are also useful in comparison of data acquired early in any given mission with data acquired many hours after takeoff.

Often overlooked is the fact that these graphs provide the analyst with a method for quick visual inspection of data. These inspections can lead to clues as to the nature of relations that may exist within and among physiological, environmental, and performance factors. With these clues, hypotheses can be formulated and tested by use of appropriate mathematical and statistical methods and models.

Rate-of-Change Analyses

The rate of change and the rate of rate of changes in physiological measures (such as heart rate) under various environmental conditions provide a sensitive index of the physiological reactivity of the astronaut to his environment. If one is interested in rates of change that occur within 10-sec and 1-min periods, the variance entry on each data sheet can be used as an index for this purpose. However, if one is interested in rate-of-change information based on periods different from those for which calculations appear on the data sheets, then the instantaneous-heart-rate raw data on each data sheet must be used since the summarizing effect of variance averages out information concerning the variability for these other periods.

In attempting to quantify rate information, it might be expected that the curve for the first and second differentials of heart-rate data would provide a measure of rate of change and rate of rate of change, respectively. However, this is not the case because of several considerations. The first consideration is that there is an assumption for the process of differentiation that requires measurements to be continuous; instantaneous heart rate is not continuous in a measurement sense. Second, since in the case of heart rate a period of perhaps 15 min is regarded as being a meaningful period, the difficulty of fitting a curve to the variations in heart rate over that period must be considered. Differential calculus does not work well under conditions of long time periods and irregularly shaped curves. Third, to quantify rate information, it is highly desirable and even necessary to arrive at a single number to represent the rate of change and the rate of rate of change of each astronaut. The question of how to get this quantification from the equation for a curve poses a difficult problem. Fourth, a desirable product would be the determination of significance of differences existing between two astronauts with respect to their rate-of-change and rate-of-rate-of-change characteristics; here again calculus does not provide a basis for the determination.

The solution seems to lie in using the concept of differentiation where it applies, and in using other techniques to modify the concept where it does not apply (ref. 14). Using table 4, the method for accomplishing this is described in the following paragraphs.

Columns 1 and 4 each contain 15 items of heart-rate data for subjects A and B, respectively, over a certain period. If one of these heart rates were calculated every 10 sec, the period would span 150 sec, or 2.5 min, for each set of data. (Raw data can be used instead of 10-sec averages

TABLE 4.—*Hypothetical Heart-Rate (HR) Data and Their Statistical Treatment*

Subject-A			Subject-B		
(1) Heart rate	(2)* $\frac{d(HR)}{dt}$	(3) $\frac{d^2(HR)}{dt^2}$	(4) Heart rate	(5)* $\frac{d(HR)}{dt}$	(6) $\frac{d^2(HR)}{dt^2}$
87			87		
	1			2	
86		0	85		1
	1			1	
85		1	84		3
	0			4	
85		1	80		1
	1			3	
84		0	83		0
	1			3	
83		1	80		1
	2			4	
81		2	84		2
	0			2	
81		0	86		1
	0			1	
81		1	85		2
	1			3	
82		1	82		1
	0			2	
82		1	80		0
	1			2	
83		1	82		0
	2			2	
81		1	80		0
	1			2	
82		0	82		0
	1			2	
81			84		
Mean, 82.9 S.D., 1.44	Mean, 0.86	Mean, 0.76	Mean, 82.9 S.D., 2.23	Mean, 2.36	Mean, 0.92

*Testing of significance of differences between average rates of change (columns 2 and 5) for subjects A and B. Mann-Whitney U-test (ref. 15): $U=19$, $P<.001$ (difference is significant).

if assessment is required for a smaller incremental rate of change.)

The mean heart rate of subject A for the 15 entries during the recorded period is 82.9 beats per minute; his variability is 1.44 beats per minute. The mean heart rate for the 15 entries for subject B is exactly the same (82.9), but the variability is different (2.23).

The question arises: Is subject A different in his rate of change of heart rate from subject B? To answer this, the first differential or (perhaps as it should be stated here) the first set of differences can be found. (While heart rate itself may be assumed to be a first differential, it is treated here as basic data.) To follow this procedure the first heart rate in a column is subtracted from the second, ignoring signs, and recording the differences as shown in column 2. The second heart rate is then subtracted from the third and recorded. Continuance of this process provides a column of numbers that, if plotted against time, yields the curve produced by the first differential, or a curve of rate of change. However, since it is desirable to avoid dealing with curves, this column of differences is simply averaged, thereby arriving at the mean difference in heart rate during the 150 sec, or the mean rate of change for the subject during this period. The signs in this computation are disregarded since the analyst is interested in only the rate of change and not the direction of this rate. In terms of calculus, one would be dealing with a nondirectional derivative. In a like manner, table 4 shows that the column of first differences, as well as its mean, also has been calculated for subject B.

The question now is: How does one arrive at a statement of the significance of difference between the mean rates of change in the heart rates of the two subjects? Usually a t-test is used in such situations, but a t-test cannot be applied in its usual form here because the variability of heart rate is related to the degree of rate of change between the two subjects. Thus, when the subjects differ as to the rate of change, the homogeneity-of-variance assumption of t cannot be met. Furthermore, there is some question as to whether the data are indeed at more than the ordinal level of measurement. Unless the interval level may be assumed, use of t cannot be justified; and in future applications of this technique, badly skewed distributions may be expected. Thus, for the general case, use of the parametric t-test is not recommended as a test of significance; instead its nonparametric counterpart, the Mann-Whitney U-test (ref. 15), should be used. Applying the U-test to the above data to determine the significance of the difference between the two mean rates of change, the absolute difference of 1.50 is found to yield a U of 19. The difference in rate of change between the two subjects is highly significant ($P<.001$).

To measure the difference between the two subjects in terms of their rate of rate of change, the second differential or set of second differences is calculated—from the column of first differences as shown in columns 3 and 6. Note that the technique for securing the second differences is exactly the same as was used for securing the column of first differences. The mean of the second-differences column is secured for each subject, and then their mean difference is tested for significance in the same manner as were the differences for rate-of-change information. The significance of difference of the rates of rate of change has not been worked out for these data as an example.

*Some Computer Programs**

Time-line-data programs—This program computes the means, variances, standard deviations, and standard scores for all incoming digital in-flight and appropriate ground-based medical data. During the Mercury missions the analog medical data had to be converted to digital form before this program could be used; an example of its type of outputs is shown in figure 2.

Distribution program—This program computes the means, variances, standard deviations, standard errors of the means, and critical ratios for skewness and kurtosis for a series of variables. The program is used mainly for assessing the shape of the distribution of selected data as indicated by the ratios for skewness and kurtosis.

The ratio of skewness indicates the degree of significance to which the particular distribution varies from the normal distribution and the direction in which it varies (positively or negatively skewed). The critical ratio of kurtosis indicates the significance of the peakedness characteristic of the distribution; that is, the extent to which the distribution is flat (platykurtic), medium-peaked (mesokurtic), or highly peaked (leptokurtic). These two characteristics are of interest due to assumptions of a normal distribution in many of the uses of the mean, variance, standard deviation, and standard error of the mean in statistical work. If the distribution is not normal, the critical ratios of skewness and kurtosis reveal this discrepancy, and a transformation of the scores is required for appropriate application of the analytical methods to the data.

Chi-square and frequency-distribution program—This program computes the means and standard deviations, and classifies data or subjects in up to seven categories of response, as well as the percentage in each category. The program also computes a seven-category as well as a three-category chi-square (ref. 16). Using the seven-category chi-square, one can test either for significant differences among up to seven astronauts with respect to one selected type of measurement under identical conditions, or for significant changes taking place in one astronaut with re-

*I thank Benjamin Fruchter, D. V. Veldman, and Earl Jennings, University of Texas, Austin, who wrote several of these computer programs.

spect to one selected type of measurement during seven different periods. Tables 5 and 6, respectively, illustrate these possibilities.

In application of this program, two separate chi-square tests are computed. First the distribution of observations in seven categories is tested for fit against the expected distribution, usually one in which the probability of the observations (data) is equal in all categories. Thus, if there is in fact no difference in the observations, each category has an equal opportunity of being of the same magnitude. The statistical hypothesis tested is the null hypothesis, and the chi-square test then determines the probability that the observed distribution of data was derived from the theoretical distribution.

Secondly the observations are classified into three categories. The computer program sums the observations (or data) in categories 1, 2, and 3; next it sums the data in category 4 (neutral category); and then it sums the data in categories 5, 6, and 7. The chi-square test is then applied to determine the significance of any difference, in observations (or data) at either end or in the middle of the distribution, from the expected distribution. The resulting chi-square is printed along with its level of significance. If the chi-square is significant at a present level of probability, one is justified in rejecting the hypothesis that the observed distribution is derived from a population with equal frequencies in the three categories.

TABLE 5.—*Seven Astronauts: One Type of Measurement Under Identical Conditions*

Seven-Category Chi-Square

Astronaut	1	2	3	4	5	6	7
Frequency	—	—	—	—	—	—	—
Percentage	—	—	—	—	—	—	—

TABLE 6.—*One Astronaut: One Type of Measurement During Seven Different Periods*

Seven-Category Chi-Square

Time	1	2	3	4	5	6	7
Frequency	—	—	—	—	—	—	—
Percentage	—	—	—	—	—	—	—

Three-way analysis-of-variance program—This program computes a three-way analysis of variance, printing out a complete analysis-of-variance summary of all combinations of all means. It permits the application of a factorial design in which subjects or data can be classified along three separate dimensions; for example, two levels of performance (two levels operationally defined), three levels of oxygen content (O_{21}, O_{22}, and O_{23}), and three time periods (first 15 min after zero-g, a 15-min sample after 1 hour of zero-g, and a 15-min sample after 2 hours of zero-g).

Analysis of variance, then, permits simultaneous comparison of data that are arranged in a particular manner and classified according to certain dimensions. All combinations of means are compared, and contributions to variance are analyzed. Differences between means and combinations of means are analyzed for significance beyond those attributed to chance probability. If differences are significant, inferences can be made for the classifications of the dimensions employed.

Correlation and regression program—This is a comprehensive computer program that computes the means and standard deviations for a series of variables, further computing the intercorrelation matrix and a complete multiple-regression analysis that provides beta weights, multiple R-squares, variance, multiple correlation, corrected multiple correlation, and R-ratio. Use of this program with in-flight medical data has been restricted to correlation. The correlation matrix provides the analyst with information as to how one measure relates to another for a given period and condition over a sample of subjects; for example, one can solve for the degree of relations among such measures as heart rate, respiration, acceleration, voice, performance, and carbon dioxide partial pressure.

Factor-analysis program—This is a comprehensive computer program that computes means, standard deviations, principal-axis-factor analysis, varimax rotation, multiple regression, factor-score weight-estimation, and standardized factor-score computation. Useful methods employed in factor analysis are explained (ref. 17), together with many examples of practical applications. Factor analysis is used for analysis of the correlation matrix to determine the common factors basic to a set of different measurements; the solution is portrayed in the form shown in table 7.

TABLE 7.—*Format of Solution by Factor Analysis*

Measure	Factor			
	I	II	III	h^2
Heart rate	—	—	—	—
Respiration	—	—	—	—
Acceleration	—	—	—	—
Blood pressure	—	—	—	—
Voice	—	—	—	—
Performance	—	—	—	—
Muscle tone	—	—	—	—
Electroencephalogram	—	—	—	—
Galvanic skin response	—	—	—	—

When conducting ground-based studies—for example, if the solution to a factor-analysis problem resulted in a high rating on factor-1 under all conditions, for both galvanic skin response (GSR) and muscle tone (EMG)—one would have evidence that GSR could be measured without measurement of EMG—one of the measurements would not be required. This type of information can be very useful and valuable in view of the current limitations on space, weight, and power within capsules during space missions.

Statistical-Model Limitations

Since many of the data represent time series, there are several inherent difficulties in their analysis. The main difficulty is that repeated observations of a measure over time are often sequentially dependent. This dependence, indicated by serial correlation, complicates the application of statistical methods that assume independence of observations.

Increase in the time interval between observations usually reduces the amount of sequential dependency. Nevertheless tests are required for determination of whether the observations (or measurements) selected for analysis are reasonably independent before statistical methods that assume independence can be applied. The principal method used in testing the serial dependency of observations is autocorrelation, which is correlation of the series with itself retarded by one or more time periods.

INTEGRATED-COMPUTER-SYSTEM CONCEPT

In addition to the time-line medical-data and associated computer programs described in this chapter, there have been significant concurrent data-analysis programs under development, aimed at having each type of data form an input into an integrated computer system (fig. 4).

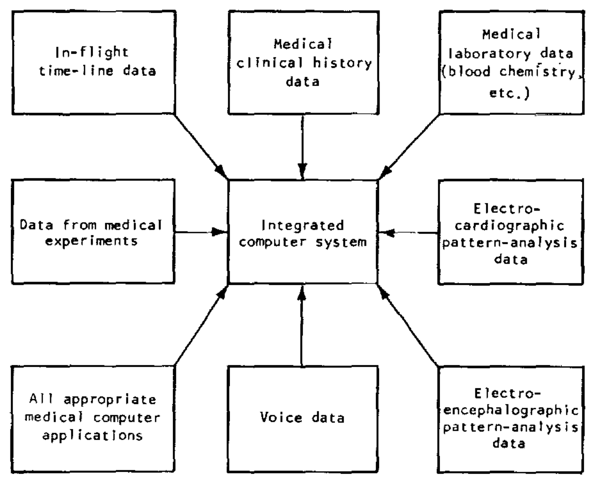


FIGURE 4.—Concept of integrated computer system.

The purpose then is to have relevant on-line data as well as appropriate stored medical data available and "on call" to the physician, through the integrated computer system, to promote maximum safety for our astronauts and to ensure the basis for proper analyses of all data relevant to future missions.

REFERENCES

1. BERRY, C. A.: Aeromedical Preparations. In Mercury Project Summary Including Results of the Fourth Manned Orbital Flight. NASA SP-45, 1963, pp. 199-209.
2. CATTERSON, A. D.; MCCUTCHEON, E. P.; MINNERS, H. A.; AND POLLARD, R. A.: Aeromedical Observations. In Mercury Project Summary Including Results of the Fourth Manned Orbital Flight. NASA SP-45, 1963, pp. 299-326.

3. Chambers, R. M.: Long-Term Acceleration and Centrifuge Simulation Studies. U.S. Naval Acceleration Laboratory, Johnsville, Pa., 1963.
4. Fraser, T. M.: Philosophy of Simulation in Man-Machine Space Mission System. Lovelace Foundation for Medical Education and Research, Albuquerque, N. Mex., 1964.
5. Lovelace, W. R., II; Schwichtenberg, A. H.; Luft, U. C.; and Secrest, R. R.: Selection and Maintenance Program for Astronauts for the NASA. Paper presented at Second International Symposium on Submarine and Space Medicine, Stockholm, Sweden, Aug. 18, 1961 (Also available in Aerospace Med., vol. 33, no. 6, June 1962).
6. Link, M. M.: Space Medicine in Project Mercury. NASA SP-4003, 1965.
7. Wilson, C. L., ed.: Project Mercury Candidate Evaluation Program. WADC TR 59-505, Wright Air Development Center, Ohio, 1959.
8. Minners, H. A.; Douglas, W. K.; Knoblock, E. C.; Graybiel, A.; and Hawkins, W. R.: Aeromedical Preparation and Results of Post-Flight Medical Examinations. *In* Mercury Results of the First U.S. Manned Orbital Space Flight. NASA, 1962, pp. 83-92.
9. Lindsey, J. F., Jr.; et al.: Evaluation of Human Factors Aspects of the B-58 Aircraft. Air Proving Ground Center TN 59-73, Eglin Air Force Base, Fla., Dec. 1959.
10. Dixon, W. J.; and Massey, F. J.: Introduction to Statistical Analysis. McGraw-Hill Book Co., Inc., 1957.
11. Edwards, A. L.: Experimental Designs in Psychological Research. Holt, Rinehart and Winston, Inc., 1960.
12. Weinberg, G. H.; and Schumaker, J. A.: Statistics, An Intuitive Approach. Wadsworth Publishing Company, Belmont, Calif., 1962.
13. Starkweather, J. A.: Variations in Vocal Behavior. U.S. Public Health Service, 1963.
14. Townsend, J. C.; and Lindsey, J. F., Jr.: Determination and Evaluation of Rate Measurements in the Analysis of Space Medical Data. Multivariate Behav. Res., vol. 2, Jan. 1967, pp. 64-70.
15. Siegel, S.: Non-Parametric Statistics. McGraw-Hill Book Co., Inc. 1956.
16. Jones, R. L.; and Lindsey, J. F., Jr.: Human Factors Aspects of Low-Altitude Flight. PGN Document 64-1, Eglin Air Force Base, Fla., 1964.
17. Fruchter, B.: Introduction to Factor Analysis. D. Van Nostrand, Princeton, N.J., 1954.

CHAPTER 4

AUTOMATED MEDICAL-MONITORING AIDS FOR SUPPORT OF OPERATIONAL FLIGHT

Robert L. Jones and Edward C. Moseley

The Apollo program implements a significant development in medical support of NASA's operational missions—the use of real-time, automated, biomedical-data-processing, monitoring techniques in flight support. Concepts of medical monitoring have evolved through the Mercury and Gemini programs, and the concomitant sophistication of hardware, software, and procedural systems, coupled with multigoaled missions of increasing complexity, has resulted in evolution of a concept of biomedical mission support requiring real-time, automated data processing. Missions have lost their flight-test emphasis and are becoming test-beds for a scientific description, analysis, and prognosis for man in space. The research mission of yesterday is becoming the operational mission of tomorrow.

Biomedical data have been provided to demonstrate man's ability to survive and perform intravehicular and extravehicular tasks involving cognitive complexity and spontaneous flexibility under conditions of severe psychophysiological and motor stress. Now our biomedical data must allow for detailed description and analysis of man as he operates in the space environment, providing empirical data for medical decisions having impact on real-time flight-planning.

With progress in the early efforts to develop studies for the improvement of management of biomedical data by use of computerized statistical techniques, and as the needs of medical support and research programs grew, the Biomedical Data Management responsibility was transferred to the Medical Research and Operations Directorate at Manned Spacecraft Center (MSC) to meet the requirement for an integrated program leading to improved operational systems for medical-data reduction and analysis. The Biomedical Data Management group is in the Biomedical Research Office, and its functions include (1) improvement in acquisition, reduction, storage, retrieval, and analysis of in-flight medical data; (2) establishment of a coordination/management focal point for biomedical data; and (3) evolution and establishment of criteria and requirement specifications for future biomedical data for subsequent and advanced manned missions, as well as for in-flight biomedical experiments.

Efforts involved in meeting these functional requirements may be better understood if they are classified as (1) for mission support of manned space flight; (2) for medical support of tests of manned systems; (3) for ground-based bioenvironmental-research programs; or (4) for support and development studies of techniques and methodologies.

Since space does not permit discussion of all four areas, our major concern will be description of the major aspects of automated medical-data processing currently planned for support of Apollo. This was a significant event in that it marked the beginning of automated preprocessing, reduction, storage, computation, retrieval, display, and analysis of bioenvironmental data for real-time flight-support.

Anyone conversant with the state of the art of medical data may well wonder why implementation of such techniques should be singled out as a significant event, in view of the many clinical and research medical institutions having well-organized, on-line, real-time, automated medical monitoring as part of their daily routines. The answer lies in the basic difference between two orientations and recalls the old specter of basic versus applied. While this issue has always been a good academic argument, it becomes very real and binding in the operational flight-testing environment when one must justify in great detail:

(1) Astronomic funding requirements to budget personnel

(2) Complex bioinstrumentation requirements to the engineers

(3) Tremendous communications requirements to the already overloaded network engineer

(4) Weight/electrical/stowage/procedural interface requirements to the spacecraft-hardware man

(5) Software requirements to the real-time programmers

(6) Long lead-time delay factors to the program manager

All this and much more is compounded by the familiar attitude of the pilot, who dislikes monitoring of his performance, sees no real need for it, and is not really convinced that you are doing anything worthwhile. The simple fact of the matter is that statements about the "search for truth and knowledge" have difficulty in standing the test as justification for the operational situation. The justification must be very definite and very real.

Of the many problems involved in such an effort this may be one of the most subtle yet outstanding. The professional bioenvironmental researcher is trained to avoid the arbitrary, disbelieve the absolute, and qualify his statements, findings, and recommendations; he deals with measures that have floating base lines, are highly individual, for which the standard error of measurement is often great, and for which the standard deviation often exceeds the mean. But when he participates in the operational situation he is competing with disciplines of opposite orientation, and these opposite viewpoints are most often in crucial management and funding organizations.

Consequently we must learn to make a philosophical, conceptual, and verbal transition from the research orientation to the operational situation if we are to participate and to carry out our ultimate responsibility. We must not be hesitant in stating our criteria and requirements. Historically when we have delayed in doing so, someone (how well qualified?) has been quick to do so for us. This is especially true in a long-lead-time, time-critical program.

Another difficulty in resolution of biomedical-data problems is achievement of optimum dialogue between the medical community and the electrical-engineering and computer communities. Many engineers have extreme difficulty in resolving the dissonance generated by first glance at biomedical problems. Floating base lines, reverse polarity, ambiguous signal-reference points, noise-filtering problems, preprocessing problems, and software problems are extensive; much time and effort are required if the engineer is to be effective in this particular area.

On the other side of the coin, the biomedical researcher often knows very little about the software/hardware/procedural/lead-time interfaces involved in his effort, and far too often he refuses to make adequate efforts to correct this lack of knowledge. This problem can be extremely critical, and the Biomedical Data Management group tries to assist by providing communications between the two communities and by becoming familiar with spacecraft hardware and procedures. Among the specific considerations emphasized in this interface are the following:

(1) Long lead-time (12 to 18 mon) for changes due to budget-cycle impact and hardware, software, and procedural changes necessary at MSC, Goddard Space Flight Center (GSFC), and the remote sites.

(2) Requirement for early (by 12 to 18 mon) specification of the following data factors: (a) variables being measured (number and kind); (b) sampling rate per variable; (c) number of channels required; (d) real-time and/or near-real-time needs (display, printouts, plots, processing, computations, etc.); (e) sampling procedures (how many subjects?; how often per man?; how long per man?); (f) required real-time identification codes, event markers, time codes, etc.; (g) real-time transmission to ground of recordings aboard spacecraft, or requirement for high-speed-dump tape recorders; and (h) priority of medical data relative to other data (because the specified pulse-coded modulation (PCM) bit stream can be used only in near-Earth orbit; circuit margins preclude use at lunar distances, so that data capabilities are severely limited).

Obviously it is often quite difficult to specify such factors in great detail as early as 18 mon before the event. However, it is well to remember that, if the lead-time is not met, an investigator may very well find himself designing hypotheses to fit the limitations of the system, and this procedure has little to offer an orderly search for knowledge.

GENERAL MEDICAL-MONITORING SYSTEMS

While use of digital computers has been widespread in research, library, administrative, and storage activities, only recently have they been used successfully in continuous monitoring of physiological data. Various historical, economic, clinical, and methodological reasons can be advanced to explain this delay in development, most of which are related to the basic nature of the analog signal (electrocardiographic, electroencephalographic, galvanic skin response, etc.) generated by some critical physiological systems.

Historically the analog computer, with its direct measurements and limited memory, preceded the digital computer with its indirect measurement by counting of numbers expressed as digits and its extensive memory. Economically the all-digital method is slow and expensive because the input data are not usable in their original form.

Clinically since Einthoven introduced the string galvanometer in 1903 (ref. 1), most medical experience has been in visual interpretation of the rhythm, patterns, rate amplitude, and duration from direct-analog strip charts. Such a scheme is basic to the current Gemini system. We might have quite a different set of clinical impressions if relative change had traditionally been viewed as opposed to the absolute measurement (ref. 2).

Methodologically repeated measurements from the same subject violate the assumption of independence necessary for use of some powerful statistical models developed for other types of applications. The number of leads, the typically high frequency of sampling, the lack of criteria, and irrelevant "noise" in the data all pose additional methodological and operational problems. In view of the possibilities of combinations of such problems, it is surprising that any real progress has been made in physiological monitoring—yet it has.

To gauge some of the progress we may focus momentarily on one physiological signal—the electrocardiographic. In general medicine we find thousands of records being collected, digitized, and summarized by cardiologists studying lead systems, noise-reduction, data-compression, work capacity, pattern-analysis, disease profiles, etc. In most of these applications the examinee is usually resting, engaged in a standard activity, or under some other controlled conditions, and use of computers in research is extensive. In general medicine and in operational use we find examination by single electrocardiogram (ECG) in the physician's office, continuous monitoring in a coronary-care unit, or magnetic-tape recordings obtained while individuals are engaged in normal activities. In these applications the examinee is generally under less-controlled conditions; the data remain in analog form for individual clinical examination; and the capital investment in hardware is quite modest.

In aviation medicine and for research purposes we find intense activity in electrocardiology. For example, in 1957 a central USAF ECG Repository was set up, and by 1959 about 67 000 normal 12-lead tracings from healthy active officers were interpreted and related to background variables, thus providing base-line data (ref. 3). Much of the current activity is in identifying and quantifying evidence, of central-nervous-system (CNS) arousal, from ECG's while the subjects are subjected to operational stresses in a simulator.

In short, highly sophisticated monitoring systems have been developed for operational ground-based use, and others have been developed for study of the general problems of continuous bioenvironmental monitoring in hostile environments. The latter activities, however, are in the developmental rather than the operational stage needed for space activity.

MSC'S MEDICAL-MONITORING AND RESEARCH SYSTEM

NASA has a highly dynamic program responding to changing national goals, radical technological advances, and changing requirements resulting from exploratory projects. It is within such a moving environment that space medicine operates at MSC.

During the Mercury program the primary emphasis was on the safety of a singly manned ballistic and short-orbital flight. Launch and reentry problems were of real engineering and medical concern, while monitoring was primarily by voice. As the Gemini program got underway, the medical system was updated to handle the two-man orbital missions of up to 14 days, as well as the extravehicular activity and some medical experiments.

The Apollo and early Apollo Applications programs increased dramatically the need for medical information, bringing substantial increases in the number of crewmen, flight complexity, mission duration, and numbers and types of preflight tests of manned systems. All these factors had significant impact on requirements of medical information. Now the system had to provide for the operational flight needs of three men in orbital and lunar flights, some as long as 30 days, and some involving free-space and lunar-surface extravehicular activity. Additionally more ground-based, preflight profile data were available for use during in-flight comparisons and evaluations, such as data gathered from the world's largest space-simulation chamber at MSC; furthermore profile data gathered during Mercury and Gemini flights were available for in-flight analysis. Obviously the multiplicity of these critical interacting requirements made the systems more sophisticated.

Planners of the Apollo Applications Program envisioned the capability of provision for sustained orbital and other extraterrestrial living and working for a flight crew of several men. The Moon is merely 240 000 miles away while Mars is a minimum of 37 million miles, so it seems safe to ignore nonlunar problems in this context. Even without the problems of distant planets the changing informational needs will be considerable for the medical and engineering communities.

Before considering an automated medical-data scheme we should describe the bioinstrumentation used in acquisition of physiological signals. However, since many of NASA's publications have described in detail the Mercury, Gemini, and Apollo bioinstrumentation systems, our discussion will be brief.

Limitation in availability of telemetry channels limits monitoring to physiological variables that are considered necessary for determination of the well-being of the flight crew. In Mercury and Gemini these real-time measures were presented in analog, with no analog-digital conversion or preprocessing. In order to provide physiological data of a more comprehensive nature for postflight analyses in depth, an in-flight tape recorder was provided for recording the measurements in analog form.

Table 1 lists the types of biomedical monitoring from spacecraft of the three projects mentioned. Figure 1 shows basic differences between the

TABLE 1.—*Types of Biomedical Monitoring from Spacecraft of the Three Projects*

Factor monitored	Mercury (1-man crew)	Gemini (2-man crew)	Apollo ($\geq$2-man crew)
ECG	Yes	A&S,[a] 320 sps [b]	A&S,[a] 200 sps [b]
Respiration	Yes	Impedance method, 40 sps; axillary ECG electrode used as sensor	Impedance-pneumograph, 40 sps
Blood pressure	Yes	Manual, squeeze-bulb, brachial-occlusive system	Mechanical, squeeze-bulb, ad libitum
Body temp	Yes	1.2 sps; oral thermistor probe; intermittent	Oral, mechanical, ad libitum
PKG [c]	None	Routed with EEG to tape recorder (experiments data)	200 sps

[a] Axial and sternal.
[b] Samples per second.
[c] Phonocardiographic.

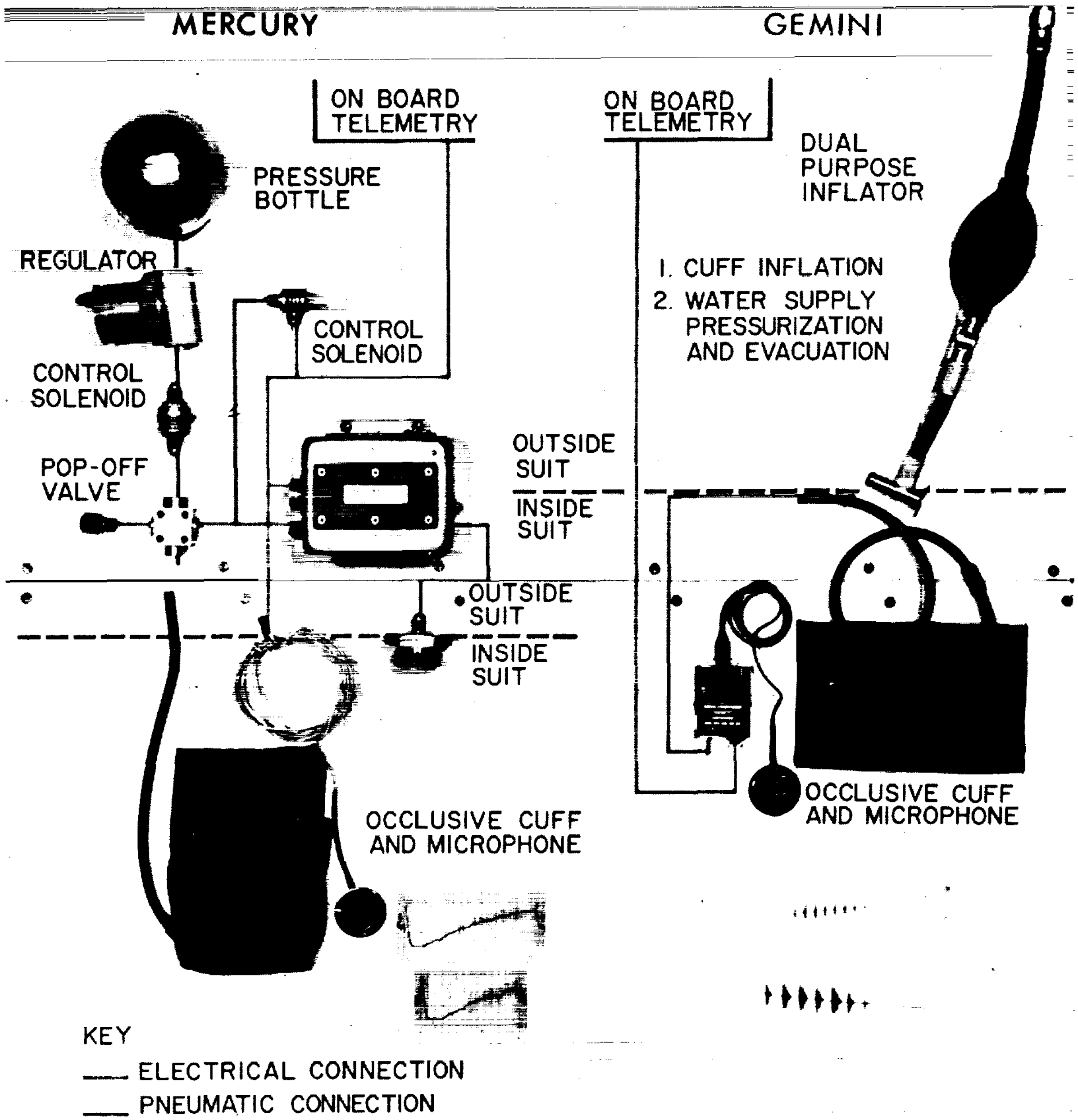


FIGURE 1.—Blood-pressure-measuring systems for Mercury and Gemini.

Mercury and Gemini blood-pressure systems. Figures 2 and 3 show the Gemini bioinstrumentation system, and figure 4 shows the Apollo system.

Voices monitored throughout all flights have furnished valuable information both during and after flights. In the Apollo program a major addition is the TV transmission that enables monitors to view the crew at selected times. A wide variety of environmental variables also are available, such as cabin pressure and temperature and suit pressure and temperature. The sampling rates for such environmental data generally range from about 0.5 to 1.2 samples per second, and the data are transmitted with PCM (pulse-code modulation).

The physiological information from Gemini

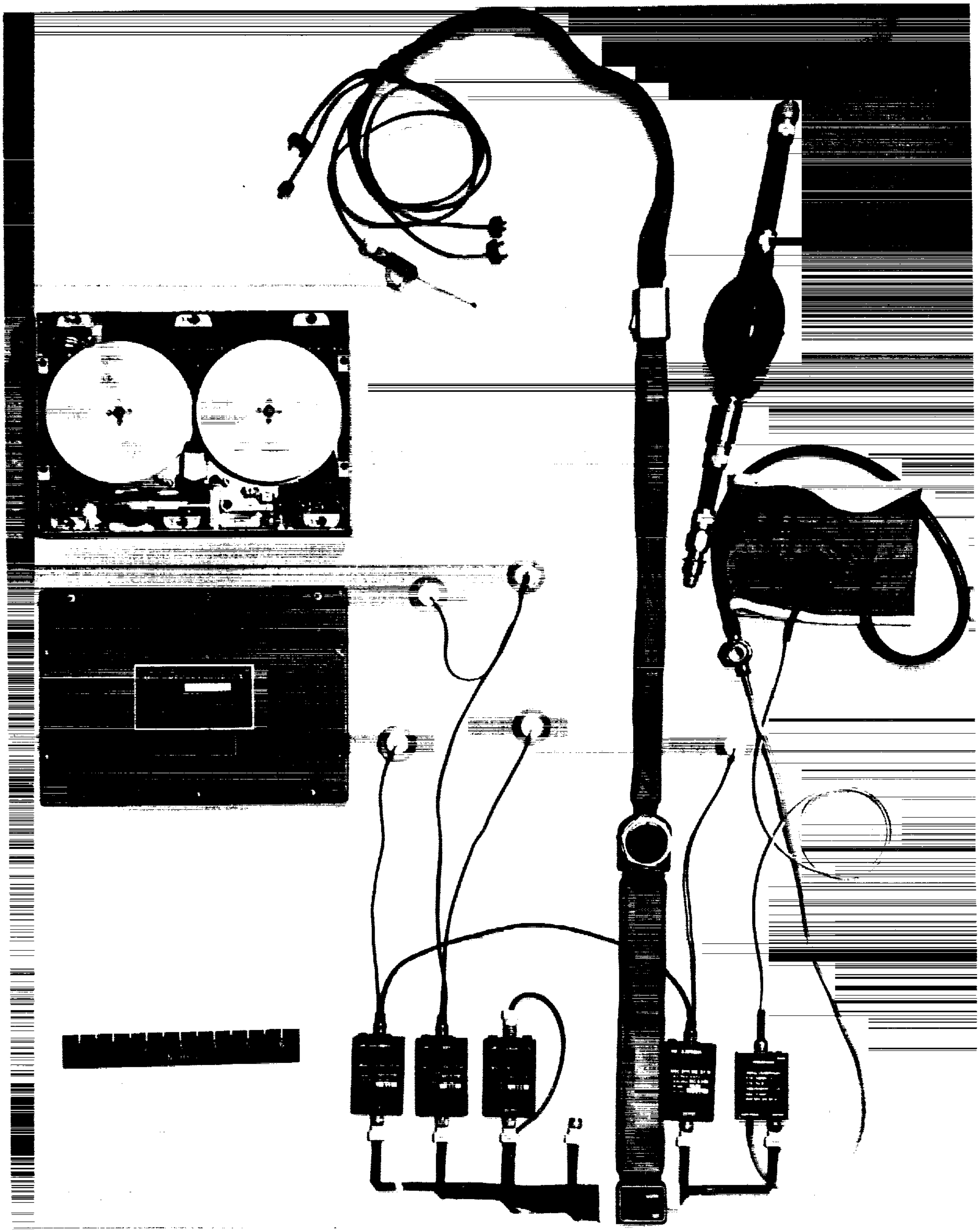


FIGURE 2.—Bioinstrumentation system for Gemini.

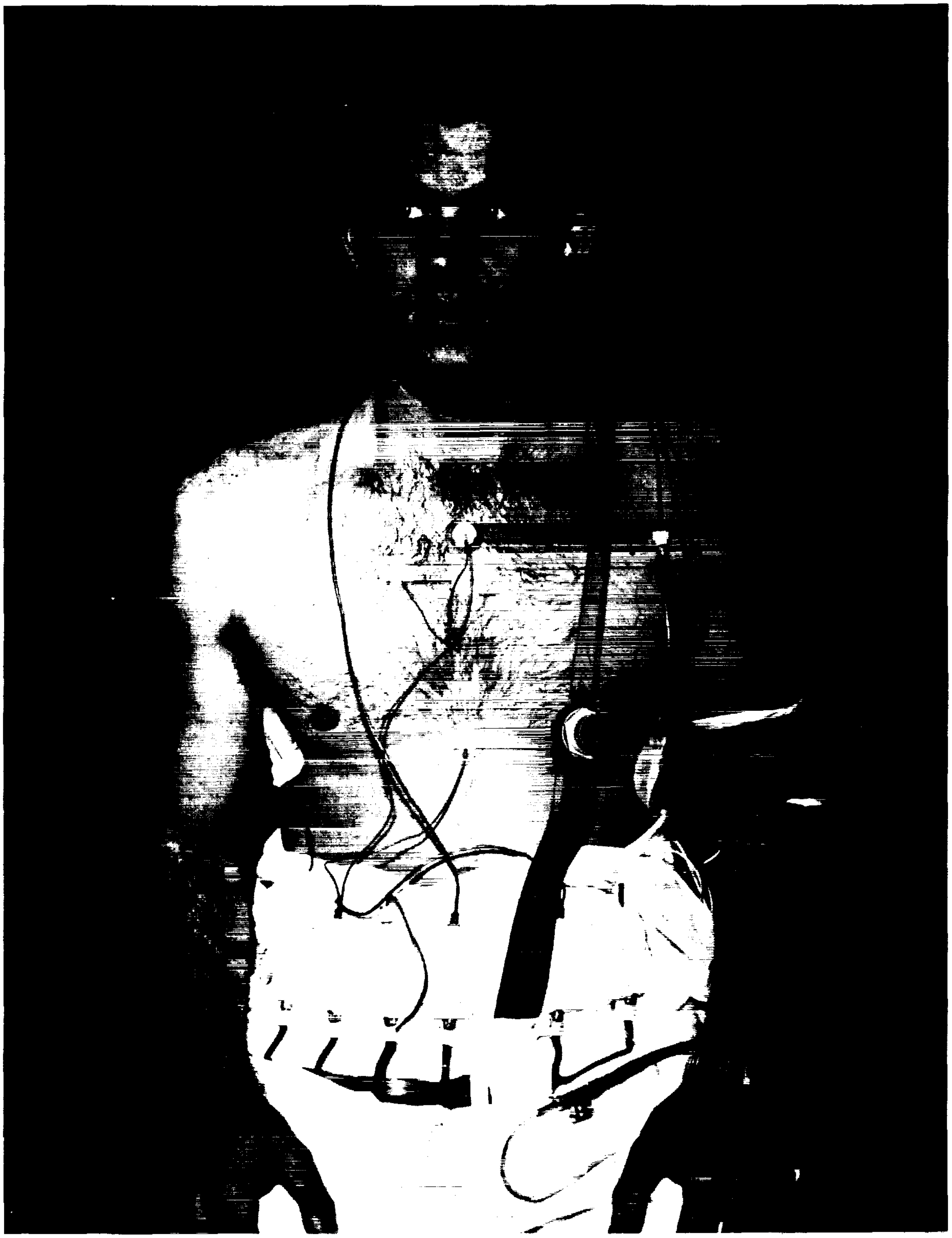


FIGURE 3.—Fitted bioinstrumentation system for Gemini.

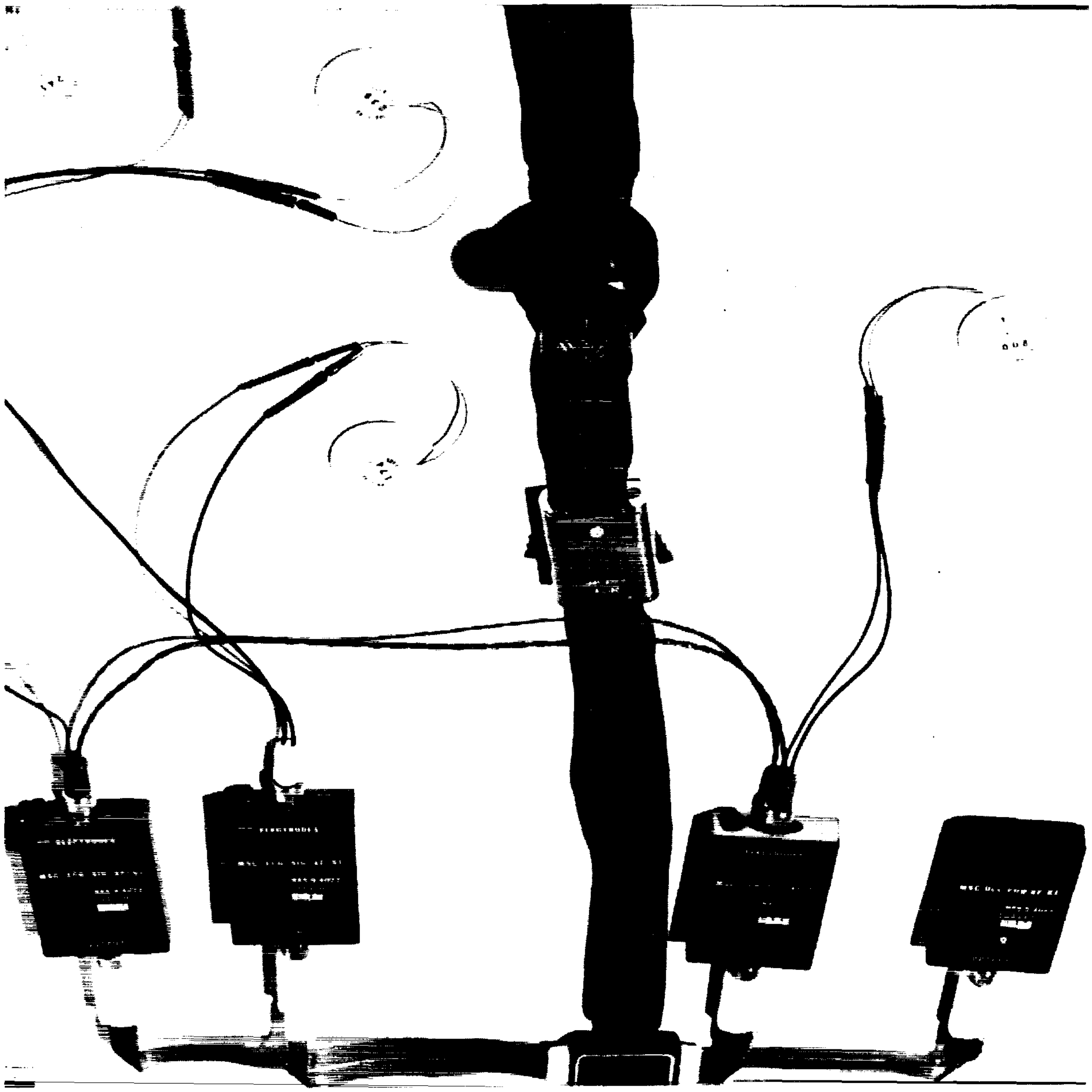


FIGURE 4.—Bioinstrumentation system for Apollo.

flowed from each astronaut to the tracking station where it was evaluated by the station surgeon and simultaneously transmitted to GSFC. Thence it was transmitted instantaneously to MSC in Houston for use by mission control. Medical data from in-flight experiments were recorded during Gemini flights on spacecraft-borne tape capable of recording up to 100 hours; this analog tape was recovered after splashdown, speeded-up to ground-elapsed time, digitized, and used as input for computer analysis of the experiment. No in-flight medical experiments are planned for Apollo.

When received at MSC the real-time telemetry information is recorded on magnetic tape for postflight analysis, and the physiological parameters are displayed in analog form on strip-chart recorders. Environmental parameters are monitored by both a computer and humans. The medical monitor can also call for a digital display of these parameters on his TV screen.

Since the primary purpose of this information is to assist the flight surgeon in clinical assessment of the crew, little real-time analysis of this physiological information is completed. Post-pass physiological summaries are calculated from the analog strip charts, as are hand plots. Immediately after the flight, selected time segments of interest are converted from analog to digital information for detailed analytic and reporting purposes.

In general the "clinical" strategy has worked well in attempts to deal with the new hostile environment for short-duration flights and few astronauts. As any program expands in terms of flight duration and/or number of crewmen, the clinical approach reaches its effective limits and more-automated schemes or aids become necessary.

We can imagine that after looking at miles of strip charts our deepest impression would recall Lincoln's comment that the thing that struck him most forcibly when he first saw Niagara Falls was "Where in the world did all the water come from!" Hardware, software, and procedural changes are under way in anticipation of the deluge of information from Apollo and its successors. Fortunately not all the expected increase will occur at once, so an effective operational system can be sequentially implemented.

EARLY APOLLO MONITORING SYSTEM

Within the general background and frame of reference thus established, we can now discuss in some detail the improved biomedical-data system approved for the Apollo program. Its general configuration is shown in figure 5.

Several cardiotachometer and pneumotachometer preprocessors, an ECG-wave-form preprocessor, a 30-sec FM tape-loop recorder, and a number of event push buttons are currently being installed for biomedical monitoring at the mission-control center. The cardiotachometer will provide visible digital readouts of both the instantaneous and selectable (6-, 10-, or 30-sec) average heart rates; it also provides selectable upper and lower limits, with a red background illuminated when the rate goes beyond limits. A 30-sec average respiration rate and a 30-sec representative ECG wave form are available from the pneumotachometer and ECG-wave-form preprocessor, respectively.

With this equipment the strip charts are needed only during critical phases of the flight and are automatically started when operationally defined limits are exceeded. In addition the hardware provides real-time analog-to-digital conversion for additional real-time descriptive and analytic manipulations. An interim system, utilizing some of this equipment, was used in simulations of the first Apollo mission. These simulations demonstrated the adequacy of the revised real-time displays for medical monitoring and determined some of the functional procedural and sampling requirements for automatic summaries of physiological in-flight information.

During the first simulated Apollo flight the monitor had a variety of mid-pass and post-pass computer summaries available to him in addition to the instantaneous information (mean, median, range, and variability) on heart rate, respiration rate, ECG-wave-form components, and selected environmental parameters, such as suit temperature and cabin temperature, for the various crewmen in real time. Time plots of these parameters were also available; up to 12 hours of these summaries could be reviewed at any given moment.

Of considerable interest are the automatic daily flight summaries of physiological and environmental parameters that were made available beginning with the interim system. As far as is practicable the instantaneous data are automatically tagged with independent variables of interest. Most tagging can be done automatically by the computer (e.g., astronaut, lead, suit on or off, day), while others (e.g., sleep, exercise, and stress periods) still require operational definitions for human push-button action.

As the data are buffered and processed for post-pass summaries the latter are routed to storage buffers containing identical tags. Thus, at the end of the day, descriptive tables and/or plots can be automatically prepared for any dependent variable by any combination of independent variables. For example, heart-rate or

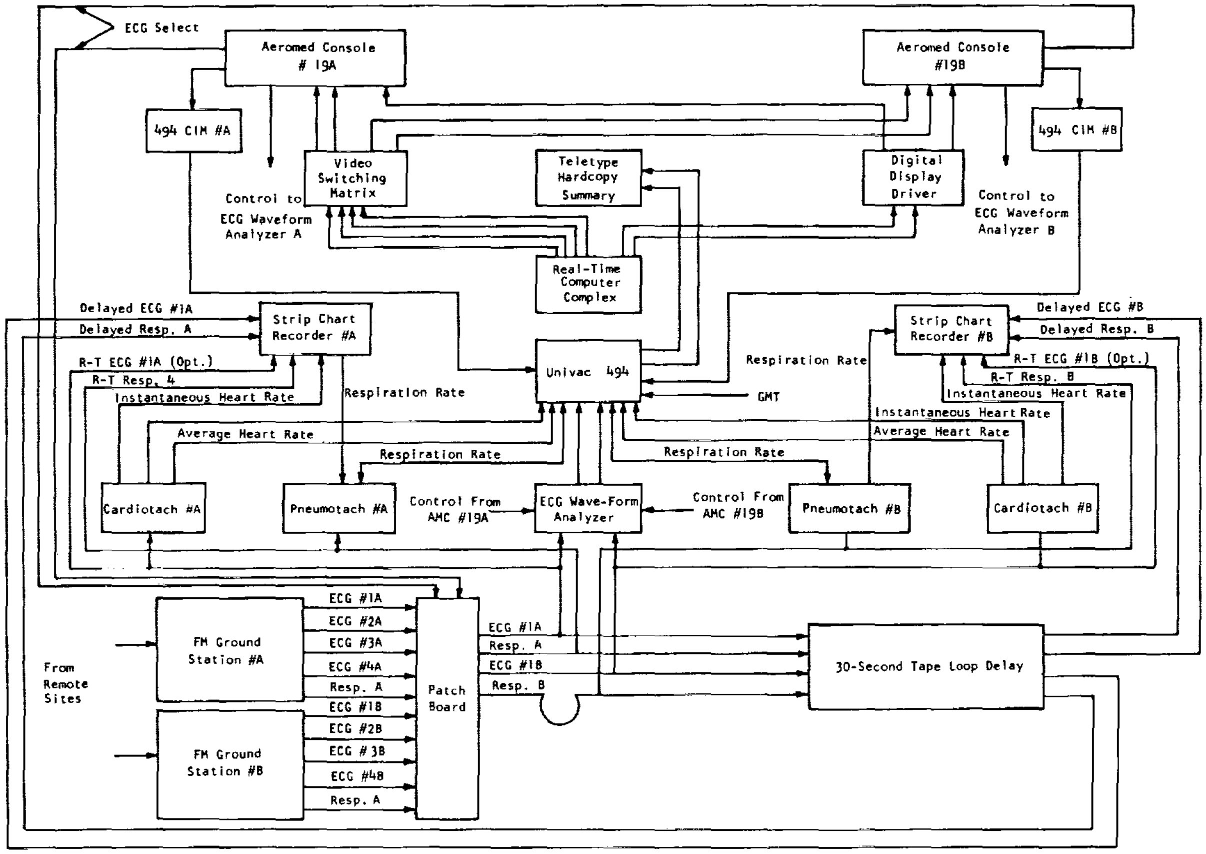


FIGURE 5.—Configuration of biomedical-data system for Apollo.

average-wave-form plots for normal activity, stressful exercise, and sleep across days to date will be available for each crewman. Furthermore, automated, annotated, Ground Elapsed Time plots can be generated daily from the post-pass summaries.

One may ask why this sort of daily summary is needed in real time when a more careful analysis could be completed after the flight. The answer, of course, is that daily trends have important implications regarding monitoring and countermeasures when the flight lasts roughly 1 mon or longer. In general the design philosophy is to do as much analog-digital conversion as possible and to strive for successive levels of real-time data-reduction.

Finally the early Apollo monitoring system will provide automatic postflight summaries; essentially they will be the daily flight summaries displayed in a format suitable for immediate reporting on the mission. They will also provide feedback for planning of the next mission, as well as forming part of the basis for a cumulative description of Apollo flights. The necessary and sufficient information for application of many inferential statistical techniques will be available in summary form and used as applicable.

IMPROVED PHYSIOLOGICAL MONITORING SYSTEM FOR APOLLO

It is difficult if not impossible to describe at this time the configuration of the future monitoring system since plans are amended daily. The higher frequency of Apollo flights, as well as their longer, overlapping, orbital, and nonorbital nature, require that the following general goals be established: (1) increased ability to plan for and detect medical crises, and (2) reduction in elapsed time between acquisition of data and formulation of interpretative reports—with indications of actions required.

For achievement of these goals, general objectives are continually being redefined in terms of the user's needs. After investigation of the physical and intellectual tools required, the practicality of each objective is determined in terms of time, usefulness, and money. Once the objectives are determined precisely and made consistent with the realities of time, relevance, and equipment, the scheduling, specifications, building, and testing are implemented. With this sort of approach we can establish attainable objectives significant to the program. Table 2 shows some idealized hypothetical requirements that may be used as criteria for evaluation of a system's effectiveness.

One objective, for example, calls for increased real-time and near-real-time displays of summary information having medical relevance. In addition to the automatic displays discussed earlier, a variety of others are now being actively considered.

Another general objective, relevant to both the planning and the detection goal, is increase in our real-time and near-real-time responsiveness to all "historical" information: data from training simulators, altitude chambers, laboratory and clinic, previous flights, and normative base-line results. Some of this information is examined before the flight and later compared with postflight results for writing of mission reports.

On the other hand, some can be fruitfully used for the making of background-comparison slides for real-time display or for some phases of "automatic monitoring"; therefore, they must be prepared as preflight summaries. Implementation at MSC of contractual efforts designed to meet these historical needs is currently being investigated.

Still another objective calls for increased use and flexibility of real-, near-real-, and non-real-time analyses. Most of the statistical analyses implied earlier are univariate descriptive statistics (e.g., mean, median, range, and variability) or univariate inferential statistics (analysis of variance, covariance, etc.). This latter class of

TABLE 2.—*A Hypothetical Systems Checklist*

I. System requirements—Does the system have

___1. Flexible input-output capability
___2. Multiple and flexible display properties
___3. Analog-to-digital conversion
___4. Digital-to-analog conversion
___5. Effective "interrupt" capabilities
___6. Efficient search and "locate"
___7. Versatile, selective extraction properties
___8. Easy proofing, updating, and correction
___9. Complete information about information
__10. Varied data-compression capability
__11. Automatic and readable report
__12. Efficient monitoring and programs
__13. Complete error messages
__14. Responsiveness to multiple users
__15. Adequate storage capabilities (tapes, core, drum, and/or disk)
__16. Easily implemented changes
__17. Adequate documentation
__18. "State of hardware" best suited to needs
__19. Adequate auxiliary equipment
__20. Priorities that permit reasonable turnaround time
__21. Varied, efficient, and useful analytical programming capability at the descriptive__, inferential__, predictive__, and interpretive levels__
__22. Adequate expansion properties
__23. Success criterion

II. Organizational requirements—Does the systems organization provide

___1. Well-defined and limited goals
___2. Coordination among information sources
___3. Personnel with time to service users or to write programs
___4. Personnel with time to improve system
___5. Interdisciplinary services
___6. Long-range administrative and monetary support

III. Operational requirements—Is the system

___1. Easy to learn for one-shot, intermittent, and repetitive applications
___2. Easy to use in one-shot, intermittent, and repetitive applications
___3. Easy to check
___4. Informative by providing timely information in usable form
___5. Interesting to the user

statistics enables us to determine, for example, whether the difference in heart rate during 3 days of sleep is simply by chance. Analysis of covariance could be used to answer the question of whether there is non-chance change in heart rate when the influence of changes in some environmental effect (e.g., temperature) is removed. Other univariate statistical models enable one to describe physiological periods, determine significant trends, eliminate noise, etc. Another class of statistics, known as multivariate, answers the same type of question but considers all variables simultaneously; for example, one can ask whether there is a non-chance difference between suit-on and suit-off conditions when all variables (i.e., rates of heart and respiration, wave-form components, temperature, CO_2, O_2, and humidity) are considered simultaneously. Other powerful multivariate models are available. Predictive statistics is still another class of statistics that can be useful when applied to some of our problems.

One should note that only the simplest descriptive statistics can be applied currently in real-time because of the workload in the Real-Time Computer Complex. On the other hand, some of the basic summary information for these methods can be generated in real time by digital computers and then applied to daily and/or postflight summaries. Still other manipulation of raw data will require analog preprocessing.

More general objectives and approaches to the future physiological monitoring system are being considered in an effort to stabilize and reduce the workload without compromising the well-being of the astronauts. The general goal is for a better descriptive, inferential, and predictive system—in about that order.

Thus the MSC biomedical-data system is indeed dynamic—by the time this paper is published, changes and improvements will have been made. However, we have described the rationale, and enhancement will come through optimization of this system within the framework of subsequent mission requirements.

REFERENCES

1. Burch, G.; and Winsor, T.: A Primer of Electrocardiography. Lea and Febiger, Philadelphia, 1945.
2. Proctor, L. D.; and Adey, W. R.: Symposium on the Analysis of CNS and Cardiovascular Data Using Computer Methods. NASA, 1965, p. 235.
3. Lamb, L. E.: First International Symposium on Cardiology in Aviation. School of Aviation Medicine, USAF, Brooks Air Force Base, Texas, 1957.

CHAPTER 5

MEDATA: A MEDICAL-INFORMATION-MANAGEMENT SYSTEM

Tate M. Minckler and Caroline L. Horton

Significant interaction between the practice of medicine and the computer sciences is inevitable. This observation is predicated on the demonstrated advantages that accrue from proper use of the computer as a tool for processing of voluminous or complex data. The interdigitation of the medical and computer sciences already has resulted in emergence of a new discipline, here dubbed medical information management, that appears destined to touch every facet of medical practice. Medical information management may be defined as a supportive activity responsible for providing coordination of informational needs and resources among the patient-care, teaching, research, and administrative services of the medical environment.

Medical information management has three functional components—record keeping, data analysis, and communication—each of which is at least partially amenable to computer support. The key to eventual integration of these now largely independent components rests upon the development of medically effective, organizationally sound, computer-based, medical record systems. For without ready access to complete data, even the most sophisticated analysis and communication facilities are severely hobbled.

Broad experience has demonstrated, however, that medical record systems do not "compute" easily. Chief among many reasons for this have been two: the medical information problem of defining specifically the content and organization of the records, and the information management problem of providing computerized support that is flexible enough to survive the constant changes in content and organization. The tasks then are to design concepts of medical documentation leading toward increased utility of the stored data, and concurrently to provide computer-supported records-management capability that neither dictates nor restricts the medical information handled by it.

An effort has been under way since July 1965 (NSR44-012-039) to provide such management capability for use with astronauts' medical information. It is with the progress toward that goal that we are now concerned.

MEDICAL INFORMATION

Medical information in this context is the body of information that is generated during patient-care activities. Its primary purpose is documentation of the factors applied to disposition of patient care. However, these records are called upon to serve research, educational, and administrative needs as well. The most compelling reason for development of a structured record system is not simply conversion to automation; rather it is provision of an accurate, detailed, and consistent bank of data. The file structure of MEDATA is designed without regard for the methods or equipment used in records management.

At least three features are essential to any systematic file structure: (1) provision for unique identification of each individual record in the file, (2) definition of the content of the file, and (3) uniform organization of the component records. We now describe these features for MEDATA applications.

Identification

The usual orientation of medical data provides a straightforward approach to record identification. Medical records are generally accumulated in terms of three parameters: the who, the what, and the when. The who may be a patient's name or any of a group of unique numbers (such as hospital number or social security number). The when is a date that may or may not be supplemented by time of day. The what defines the content in terms of the function and the type or source of the data. The content of the file is the axis on which any records system turns.

Content

In determining a pattern for medical record file structure, data function has been the paramount consideration. When the content of patient-care information is viewed from the standpoint of functional documentation, three rather distinct classes of medical reports may be discerned: the survey, the specific problem workup, and the status report. The first two of these are commonly intermixed in current practice, but can and should be separated for a variety of reasons.

Whether the survey or screening examination is applied to a population or an individual patient, its purpose is the same; it identifies from the whole the parts having a high probability of disease and thereby establishes a set of problems requiring further medical investigation. Conversely but equally significant, the survey report should also serve to document the parts that are free of recognizable disease. A complete or partial survey may be conducted at any point in the patient-care cycle, either as an independent periodic examination or to complement a specific workup.

The workup is usually initiated either because the patient seeks help for a problem or because a survey has elicited one or more potential problems; it is an attempt to establish a diagnosis and to institute appropriate therapy for each problem. The principal functional results of a workup are two: a plan of action that may be either diagnostic or therapeutic, and initial implementation of that plan (a set of orders).

The effectiveness of the plan is evaluated and documented periodically. The doctor's "progress note" is a prime example of such a status report, but also included are other reports that serve the same basic purpose, such as nurses' notes and reports of procedures.

This philosophic approach provides a framework for understanding and categorizing of the basic functions of each informational component in the patient-care environment. Each functional division (survey, workup, status) has a natural group of subdivisions based on the type or source of data; for example, the several types of survey documentation include electrocardiographic, laboratory, X-ray, and review of systems. Table 1 shows the basic pattern of record identification, listing the specific types of survey reports currently in use; each, representing a unique step in the survey process, may be generated by a different person at a separate time or in a special location, and usually reports on a different subject. These same criteria are applied to the designation of types of reports throughout the MEDATA file system.

In summary, unique identification is provided by requirement of a standardized set of statements regarding the "who," "what," and "when" of each report in the file. In the MEDATA application the patient's identity is established by the social security number (SS NO); content is defined by both the function of the RECORD and the TYPE of record, and DATE indicates the time frame.

Organization

The MEDATA system divides each medical report into two segments: the "leaders" or identification portion (already described) and the "body." An orderly approach is followed in establishment of the body of a particular report.

TABLE 1.—*MEDATA Identification Pattern for Survey Records*

Category	Heading	Date (examples)
Who	SS No.	123-45-6789
What (A)	Record	Survey
What (B)	Type	Identification Review of systems (ROS) Dental Laboratory X-ray ECG Measurement Vision Summary
When	Date	01 Jan 68

First the informational content must be defined. Figures 1 and 2 are reproductions of the Standard Report of Medical Examination (Form-88), one of the survey documents in use at Manned Spacecraft Center. Its content was adopted as the starting point for development of the MEDATA system.

Terms must then be chosen as headings to represent the contents. Unidentified data obviously are meaningless. Although many data when presented in context need no overt headings, the heading concept is implied by the context. For this reason, most items of data in any file are accumulated as direct answers to the questions posed by headings. This sequential association of heading followed by data is called a heading/data pair.

The headings must be organized into an outline so that relations are clear among the various blocks of data. As an example of this procedure, consider items 57 and 58 of Form 88 (fig. 2), measurements of the cardiovascular system. Figure 3 is an outline arrangement of these basic items, with minor modifications in terminology but with all original content intact; notice the use of standard outline techniques to relate ideas. All data about pulses are indented under that term, and other degrees of indentation clearly establish the relations among the remaining headings. This hierarchy (outline) technique provides a pattern by which man (and machine) can recognize the organization of headings and therefore the association of data.

The final consideration in establishment of a systematic file structure is the policy regarding data "formatting." The format of an item of data refers to the detailed characterization of the way in which that item is to be reported. The MEDATA system recognizes three fundamental and distinct kinds of information that must be reported in a medical context: quantities, judgments, and facts.

Quantities are measurements generated in the medical environment directly or indirectly from the patient; height, weight, blood pressure, and laboratory values are obvious examples. Quantities are formatted by making the numeric value the first element in the data area or field; this is generally followed by the type of unit in which the value is reported—height, 72 in.; or Hb, 15.5 g.

Judgments are items for which the examiner is expected to evaluate one or more criteria and summarize his opinion. With the MEDATA system this expression is formatted by entering of POS (positive) or NEG (negative) followed by an explanation or amplification in prose as appropriate. The entry POS means that in the opinion of the examiner the item under consideration is unusual, abnormal, or otherwise noteworthy; NEG indicates normal, within acceptable limits, or not significant. The entry POS is almost always followed by a description of the abnormal finding or some comment to explain why that item is notable; NEG may be amplified by prose as necessary. Table 2 shows a portion of the record of a physical examination. This policy permits the rapid scanning of reports for information judged to be medically significant by the examiner.

The judgment format may also be combined with quantitative reporting: the judgment (POS or NEG), plus commentary, is entered after the quantity and its type of unit—weight, 185 lb; POS (30-lb overweight for height).

Facts, the third kind of data, are not subject to quantitation or judgment in the usual sense. They are items of information, such as name, address, history of measles, etc., often derived from the patient or another informant. Facts are simply placed in the report as communicated; a judgment may be made as to the reliability of the informant, but the items themselves are not subject to medical evaluation or quantitation. The formatting of facts is not restrictive but must be uniform, item by item; for example, the question regarding name should always be answered in the same pattern—last name, first name, middle initial.

Portions of the MEDATA formats for Form-88 data will be used throughout the remainder of this presentation as specific examples of the file structure currently implemented. Forms for more elaborate capture of information are being developed.

INFORMATION MANAGEMENT

Information management, as applied to a record system, means the provision of smooth and integrated mechanisms to collect, process, and retrieve records or parts of records. For these purposes the advantages of the modern digital

Standard Form 88
(Rev. June 1956)
Bureau of the Budget
Circular A-32 (Rev.)

REPORT OF MEDICAL EXAMINATION

88-106

1. LAST NAME—FIRST NAME—MIDDLE NAME	2. GRADE AND COMPONENT OR POSITION	3. IDENTIFICATION NO.
4. HOME ADDRESS (*Number, street or RFD, city or town, zone and State*)	5. PURPOSE OF EXAMINATION	6. DATE OF EXAMINATION
7. SEX / 8. RACE / 9. TOTAL YEARS GOVERNMENT SERVICE: MILITARY, CIVILIAN	10. AGENCY	11. ORGANIZATION UNIT
12. DATE OF BIRTH / 13. PLACE OF BIRTH	14. NAME, RELATIONSHIP, AND ADDRESS OF NEXT OF KIN	
15. EXAMINING FACILITY OR EXAMINER, AND ADDRESS	16. OTHER INFORMATION	
17. RATING OR SPECIALTY	TIME IN THIS CAPACITY (*Total*)	LAST SIX MONTHS

NOR-MAL	CLINICAL EVALUATION (Check each item in appropriate column; enter "NE" if not evaluated.)	ABNOR-MAL
	18. HEAD, FACE, NECK, AND SCALP	
	19. NOSE	
	20. SINUSES	
	21. MOUTH AND THROAT	
	22. EARS—GENERAL (*Int. & ext. canals*) (*Auditory acuity under items 70 and 71*)	
	23. DRUMS (*Perforation*)	
	24. EYES—GENERAL (*Visual acuity and refraction under items 59, 60 and 67*)	
	25. OPHTHALMOSCOPIC	
	26. PUPILS (*Equality and reaction*)	
	27. OCULAR MOTILITY (*Associated parallel movements, nystagmus*)	
	28. LUNGS AND CHEST (*Include breasts*)	
	29. HEART (*Thrust, size, rhythm, sounds*)	
	30. VASCULAR SYSTEM (*Varicosities, etc.*)	
	31. ABDOMEN AND VISCERA (*Include hernia*)	
	32. ANUS AND RECTUM (*Hemorrhoids, fistulae*) (*Prostate, if indicated*)	
	33. ENDOCRINE SYSTEM	
	34. G-U SYSTEM	
	35. UPPER EXTREMITIES (*Strength, range of motion*)	
	36. FEET	
	37. LOWER EXTREMITIES (*Except feet*) (*Strength, range of motion*)	
	38. SPINE, OTHER MUSCULOSKELETAL	
	39. IDENTIFYING BODY MARKS, SCARS, TATTOOS	
	40. SKIN, LYMPHATICS	
	41. NEUROLOGIC (*Equilibrium tests under item 72*)	
	42. PSYCHIATRIC (*Specify any personality deviation*)	
	43. PELVIC (*Females only*) (*Check how done*) ☐ VAGINAL ☐ RECTAL	

NOTES. (*Describe every abnormality in detail. Enter pertinent item number before each comment. Continue in item 73 and use additional sheets if necessary.*)

(*Continue in item 73*)

44. DENTAL (*Place appropriate symbols above or below number of upper and lower teeth, respectively.*)

O—*Restorable teeth*
/—*Nonrestorable teeth*
X—*Missing teeth*
XXX—*Replaced by dentures*
(6 X 8)—*Fixed bridge, brackets to include abutments*

RIGHT	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	LEFT
	32	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	

REMARKS AND ADDITIONAL DENTAL DEFECTS AND DISEASES

LABORATORY FINDINGS

45. URINALYSIS: A. SPECIFIC GRAVITY			46. CHEST X-RAY (*Place, date, film number and result*)
B. ALBUMIN	D. MICROSCOPIC		
C. SUGAR			
47. SEROLOGY (*Specify test used and result*)	48. EKG	49. BLOOD TYPE AND RH FACTOR	50. OTHER TESTS

FIGURE 1.—Report of medical examination (Form-88), front.

MEASUREMENTS AND OTHER FINDINGS

51. HEIGHT	52. WEIGHT	53. COLOR HAIR	54. COLOR EYES	55. BUILD: (*Check one*)	SLENDER	MEDIUM	HEAVY	OBESE	56. TEMPERATURE

57. BLOOD PRESSURE (*Arm at heart level*)						58. PULSE (*Arm at heart level*)				
A. SITTING	SYS.	B. RECUMBENT	SYS.	C. STANDING (*3 min.*)	SYS.	A. SITTING	B. AFTER EXERCISE	C. 2 MIN. AFTER	D. RECUMBENT	E. AFTER STANDING 3 MIN.
	DIAS.		DIAS.		DIAS.					

59. DISTANT VISION		60. REFRACTION			61. NEAR VISION	
RIGHT 20/	CORR. TO 20/	BY	S.	OX	CORR. TO	BY
LEFT 20/	CORR. TO 20/	BY	S.	OX	CORR. TO	BY

62. HETEROPHORIA (*Specify distance*)

ES°	EX°	R. H.	L. H.	PRISM DIV.	PRISM CONV.	PC	PD
					CT		

63. ACCOMMODATION		64. COLOR VISION (*Test used and result*)	65. DEPTH PERCEPTION (*Test used and score*)	UNCORRECTED
RIGHT	LEFT			CORRECTED
66. FIELD OF VISION		67. NIGHT VISION (*Test used and score*)	68. RED LENS TEST	69. INTRAOCULAR TENSION

70. HEARING		71. AUDIOMETER									72. PSYCHOLOGICAL AND PSYCHOMOTOR (*Tests used and score*)
RIGHT WV /15	SV /15		250 256	500 512	1000 1024	2000 2048	3000 2896	4000 4096	6000 6144	8000 8192	
LEFT WV /15	SV /15	RIGHT									
		LEFT									

73. NOTES (*Continued*) AND SIGNIFICANT OR INTERVAL HISTORY

(*Use additional sheets if necessary*)

74. SUMMARY OF DEFECTS AND DIAGNOSES (*List diagnoses with item numbers*)

75. RECOMMENDATIONS—FURTHER SPECIALIST EXAMINATIONS INDICATED (*Specify*)	76. A. PHYSICAL PROFILE					
	P	U	L	H	E	S
77. EXAMINEE (*Check*) A. ☐ IS QUALIFIED FOR B. ☐ IS NOT QUALIFIED FOR	B. PHYSICAL CATEGORY					
78. IF NOT QUALIFIED, LIST DISQUALIFYING DEFECTS BY ITEM NUMBER	A	B	C	E		

79. TYPED OR PRINTED NAME OF PHYSICIAN	SIGNATURE	
80. TYPED OR PRINTED NAME OF PHYSICIAN	SIGNATURE	
81. TYPED OR PRINTED NAME OF DENTIST OR PHYSICIAN (*Indicate which*)	SIGNATURE	
82. TYPED OR PRINTED NAME OF REVIEWING OFFICER OR APPROVING AUTHORITY	SIGNATURE	NUMBER OF ATTACHED SHEETS

FIGURE 2.—Report of medical examination (Form-88), back.

```
CARDIOVASCULAR
 BP
  SITTING
   SYS:
   DIAS:
  RECUMBENT
   SYS:
   DIAS:
  STANDING
   SYS:
   DIAS:
 PULSE
  SITTING:
  RECUMBENT:
  STANDING:
   EXERCISE
    IMMED AFTER:
    2 MIN AFTER:
```

FIGURE 3.—Example of a MEDATA outline derived from the cardiovascular-measurement section of Form–88 (fig. 2).

computer and associated supporting equipment are obvious and need no further justification. This section describes a unique concept in computer programming and then covers the operational implementation of that concept for acquisition, processing, retrieval, and maintenance of MEDATA information.

The Primary/Peripheral Programming Concept

The primary/peripheral concept represents a significant departure from classical programming approaches. It offers great flexibility in file structure to the user who is not familiar with computers, and yet provides a core for smooth integration within the acquisition, processing, and retrieval components of the management system. Understanding of primary/peripheral programming requires background in the broad responsibilities of a classical computer program in relation to the job it is to perform.

The term computer program as routinely used can be loosely defined as a discrete set of instructions which, when applied to the computer, controls in explicit detail the manipulations performed by the computer system on a specified file of data. The usual computer program functions at two levels. The primary level is the logical job for which the program was initially intended (e.g., to compute standard deviations or to retrieve an existing item from the file). The secondary level of programming refers to instructions required to relate (1) data files to primary program steps, and (2) the results of primary programs to man. In the FORTRAN programming language, for example, secondary-level programming is represented by FORMAT STATEMENTS.

TABLE 2.—*Portion of the Record of a Physical Examination*

Item	Judgment	Explanation
Abdomen	NEG	Scaphoid; good tone
Anus and rectum	POS	Asymptomatic large hemorrhoids
GU	NEG	

The operational difference between primary and secondary levels becomes clear when one considers that primary programming deals with concepts (standard deviation, sort, retrieve, etc.) or generalized operations performed on many different data files. Secondary programming, quite separately, is concerned with the intimate details of a specific file (the location and arrangement of the numbers from which a standard deviation is to be calculated, or the particular items of data to be sorted or retrieved).

Secondary programming, therefore, recognizes the formatting and the sequential relations of a specific data file; i.e., it defines content and organization. Since these definitions are already contained within the basic structure of each medical-information document already described, their preservation during the acquisition of records in machine-readable form provides an automatic peripheral mechanism for accomplishing most of the functions of secondary programming.

Primary programming, then, is the concept of limiting the responsibility of a set of computer programs to performance of basic conceptual tasks. Peripheral programming is the inclusion, within the body of each computer-stored record, of the necessary secondary-level definitions of content and organization in context with the specific data. Peripheral programming obviates the need for special secondary-level program segments attached to each primary program; instead the primary program requires only the addition of a standard interface segment that can interpret the peripherally programmed definitions contained in any record. The operational details of this concept are now described under "facsimile storage."

Facsimile Storage

Facsimile storage (FACS) is our name for the construction of the computer record; it is based on preparation by use of a standard typewriter keyboard of a human-readable document in machine-readable format. Facsimile storage preserves the complete content of each individual record in the computer, including not only prosaic (English) headings and associated text of data but also all organizational relations.

The computer handles FACS in a simple manner depending only on three position codes for each heading/data pair. The first, or hierarchical code set, has two functions: it identifies the beginning of a heading/data pair, and it specifies the hierarchical relation between its own heading and the preceding and following headings. The set can be any series of typewriter symbols; this presentation uses zero through 9 (0, 1, . . ., 9). Each digit indicates progressively the degree of indentation from the left margin in the defined outline organization of headings: zero means no indentation, 1 is the first level, 2 is the second level, and so on. Table 3 shows application of the hierarchical code to the outline of headings given in figure 3.

The second code is required to separate the variable-length heading field from the variable-length data field of the heading/data pair. For this purpose the typed colon (:) serves the dual function of being understood by both man and machine. The third code terminates each heading/data pair and can be any unique character code. The present terminating code is a special unprintable character automatically generated by the system during data acquisition.

TABLE 3.—*Application of the Hierarchical Code Set to Some Headings of Form–88 (fig. 3)*

Code number	Application to heading
0	CARDIOVASCULAR
1	BP
2	SITTING
3	SYS:
3	DIAS:
2	RECUMBENT
3	SYS:
3	. DIAS:
2	STANDING
3	SYS:
3	DIAS:
1	PULSE
2	SITTING:
2	RECUMBENT:
2	STANDING:
3	EXERCISE
4	IMMED AFTER:
4	2 MIN AFTER:

In addition to the heading/data codes, editorial codes are generated at the time of document preparation and stored in the record; they control layout involving carriage return, tabulation, line feed, etc. By a simple program these codes can be translated into functional equivalents in the computer: for example, "carriage return" equals "new print line," and "tabulate" equals "begin print in [specified] space." The term facsimile storage therefore means just that—an exact facsimile of an original document resides in the computer.

The mechanisms by which peripheral programming is implemented during data acquisition are discussed next.

Systematized Terminal-Acquisition Technique

The systematized terminal-acquisition technique (STAT) is a one-step procedure for transcribing information into machine-readable language; it is user-oriented, being accomplished by the user in his own environment and under his complete control. This technique replaces the classical key-punch approach to data transcription. Using a simple typing operation, it takes advantage of the familiarity with specific medical phraseology, spelling, and abbreviations that is available only among the users' staffs.

The current implementation of STAT is designed to function in environments where multiterminal time-sharing is not yet practical for routine operations. However, the basic principles are equally applicable to on-line and off-line use.

At present, STAT functions through a commercially available terminal (IBM–1050 Tele-Communications System) equipped with a standard keyboard/printer (Selectric typewriter) as well as components that punch and read paper tape and cards. Four characteristics of this terminal make it particularly suitable for the purposes at hand:

(1) Operation of the terminal is simple. Clerical and secretarial personnel, with no background in computer sciences, can be trained in a matter of hours to perform all required manipulations.

(2) The terminal can be "programmed"; that is, instructions such as carriage return, tabulate, tape punch "on" or "off," reader stop, etc., may be punched and stored on paper tape in context with alphabetic and numeric characters. As this tape is later read through the terminal, each instruction is performed so that a degree of format control is provided that is limited only by the sequential nature of paper tape. Under program control, data may be accepted from any input device (keyboard, paper-tape reader, or card reader) and transferred to any or all output devices (printer, paper-tape punch, card punch, etc.).

(3) Human-readable and machine-readable documents can be produced automatically and simultaneously so that data are available for immediate use, even though computer support may be delayed or periodic.

(4) This terminal can function on one or both of two channels or "loops." The "home loop" synchronizes the various components attached to a specific terminal. The "line loop" adds a telephone line to the circuit, over which data may be sent to or received from another compatible terminal or computer. This two-channel feature is important. Documents can be drafted with the machine on the "home" channel; additions, corrections, or deletions can be manually entered from the keyboard after editing. The edited document (stored on paper tape) is then transmitted over the "line" channel to another receiving terminal or computer.

The direct transcription of medical information into computer-readable form (peripheral programming) is readily accomplished by medical secretaries using the terminal features just described.

Operation of STAT—The systematized terminal-acquisition technique is summarized in figure 4. The typist, sitting at the terminal keyboard, places a master tape in the paper-tape reader; this tape controls the sequential typing of a format (an organized set of headings) to which the typist adds data from a collection form. Two documents result: one is the typed page; the other, a paper tape called the data tape. Each contains an amalgam of predefined headings from the master tape and data from the collection form.

Preparation of master tapes—Each heading outline (described under Medical Information) is

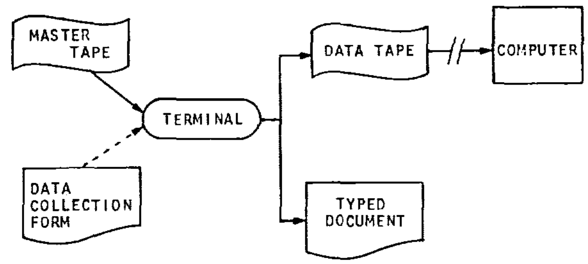


FIGURE 4.—Acquisition mechanics.

"programmed" as a master tape. Recall that the facsimile-storage technique of peripheral programming requires the use of typewriter symbols to represent indentation of the headings in the outline for retention of relations during computer storage. In this application, zero equals no space to the left of the heading, 1 equals one space, and so on through 10 levels. The outline is typed with the appropriate symbol inserted before each heading. (Editing codes, such as carriage return or tabulate, may be inserted to provide an organized appearance of the document.) A typed colon (:) separates the heading field from the data field. Wherever manual data are to be entered during operation of the master tape, a reader-stop code is punched. A special (in our case, unprintable) code is used to indicate the end of each data field. The finished master tape requires meticulous preparation but provides rapid and accurate reproduction of the original outline, to which a typist can add data at great speed.

Total operational efficiency of STAT can be significantly increased by use of the punch-card capability of the terminal system. By prepunching of cards containing the necessary instructions, the time for preparation of each master tape can be reduced from several hours to 15 to 20 min (fig. 5). A card is prepared for each of the sequential levels of a hierarchical outline (currently 10 levels). These 10 cards contain all the instructions required by the master tape except the specific words of the headings. The STAT system uses the simple method of marking each card with a number corresponding to the sequential degree of indentation for its heading. The original outline is then coded by assessment of the degree of indentation and placement of that number in the margin of the outline. A series of heading cards

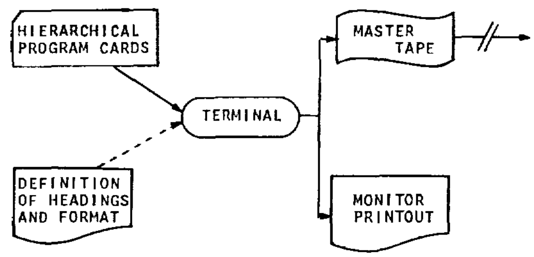


FIGURE 5.—Preparation mechanism for master tape.

```
                          MEDATA FILE

RECORD:  SURVEY
TYPE:    DENTAL

SS NO:  [redacted]               DATE:   20Jan68
                                 NAME:   Adams, John

STATUS:
 X=MISSING   O=RESTORABLE   /=NONRESTORABLE   _=PROTHESIS
 (7 x 9)=FIXED BRIDGE (BRACKETS INCLUDE ABUTMENTS) (CODED BELOW NO)

 RIGHT:     01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16   LEFT
  UPPER:     X                 (  X  X  )                  X

 RIGHT:     32 31 30 29 28 27 26 25 24 23 22 21 20 19 18 17   LEFT
  LOWER:     X        O                                    X

COMMENTS:                  Good dental function

DENTAL CLASS:              2
                           Type 3

EXAMINER:                  J B Moreton, Capt, USAF DC

TYPED:                     jlb/01Feb68
```

FIGURE 7.—Appearance of final typed document, a facsimile of which is submitted to the computer as the data tape.

is ordered sequentially according to the numbers of the outline format. This deck of cards, when read through the machine, stops at the correct indentation for the manual entry of each heading in turn. The resultant paper tape is a master tape of the outline.

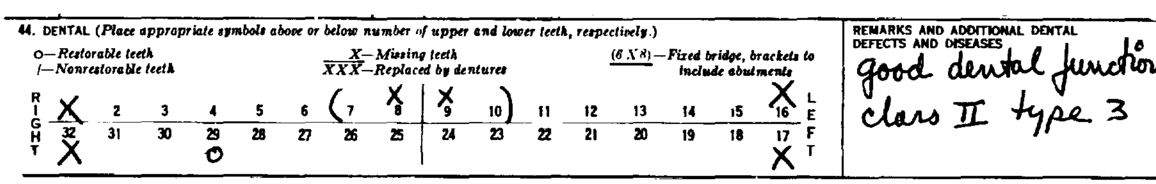
The data-collection form—One rule governs the design of a data-collection form: it must have the same headings as the corresponding master tape, or their equivalents, and similar sequential organization. It should be noted that data collection may be accomplished with this system by dictation, either directly or through a recording device, as easily as by use of a written collection form. Again, however, the sequence of the headings is critical for efficient transcription. Figure 6 shows a portion of Form–88 used as a manual-collection document and ready for transcription.

The results—Figure 7 shows the typed result from STAT. Corrections may be made by the secretary either when errors occur or during a tape-to-tape duplication of the original data tape. Editing is accomplished by proofing of the typed document, since it and the data tape have exactly the same content. When correct, information from the data tape is fed to the computer; this step may be accomplished by reading of the data tape by a paper-tape reader attached directly to the computer, or by transmission of the data over a telephone line from the paper-tape reader of the terminal.

The Information-Management Package

The computer manipulations necessary to service the MEDATA record files are provided in a collection of computer programs called the information-management package (IMP) and designed and developed on a medium-size, second-generation computer (an SDS–930 with 16-K core, three tape drives, a line printer, and a console typewriter). All programming is written in a subset of FORTRAN common to most compilers. The IMP is ultimately intended for a time-shared, multiterminal environment but is equally suitable for limited operational implementation that requires less sophistication.

Pilot projects successfully used the described STAT but without telephonic transmission to the computer. Data tapes were read by a paper-tape reader on the computer, and retrieval was accomplished by use of the console typewriter, paper tape, or punched cards to define retrieval parameters. The present version of IMP operates


44. DENTAL (*Place appropriate symbols above or below number of upper and lower teeth, respectively.*)

o—*Restorable teeth*
/—*Nonrestorable teeth*
X—*Missing teeth*
XXX—*Replaced by dentures*
(6 X 8)—*Fixed bridge, brackets to include abutments*

RIGHT	X						(	X	X	)						X	LEFT
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
	32	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	
	X			O												X	

REMARKS AND ADDITIONAL DENTAL DEFECTS AND DISEASES

good dental function
class II type 3

FIGURE 6.—Portion of Form–88 used as collection form and ready for transcription.

only in the area of record-management, but communications and analytic capabilities are planned as modules for the future.

Modularity is in fact the key to the IMP concept. The IMP is organized into two types of computer-program components: a basic control routine (monitor), and clusters of subroutines (fig. 8).

The monitor has three planned segments. The executive portion controls remote-terminal access to the system and assigns responsibility for particular actions to one of the other two segments. The executive bridge of the monitor cannot be completed until either a computer is totally dedicated to the MEDATA project, or a multiprogramming, time-shared computer system is available. This executive function is handled manually in the present version. Both limbs of the monitor are currently functional however. The "update" segment handles all modifications of the record files, including additions, deletions, and changes (corrections), as well as the sorting and storing of data. The "retrieve" segment is concerned with recall of information from the files. Both segments of the monitor act through selective use of task-specific subroutines; each subroutine performs a single, unique function and, during the time it is in use, is completely independent of the rest of the system. This approach permits the alteration, exchange, or addition of subroutines with minor or no changes in the basic monitor program.

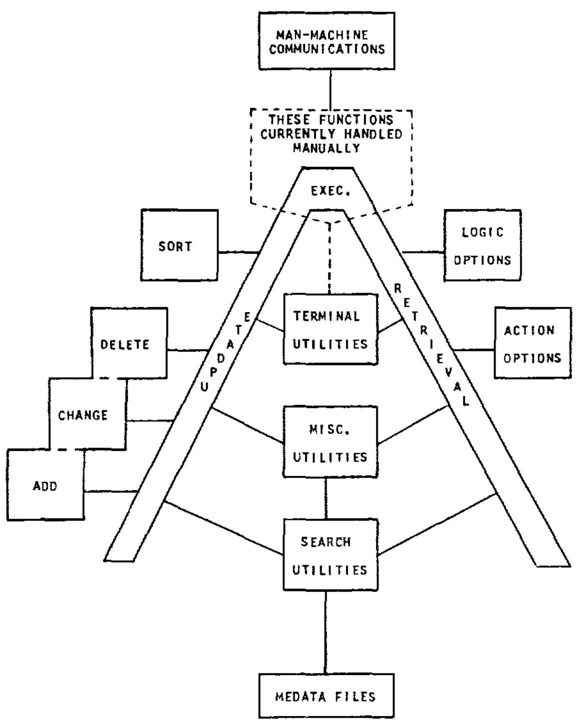


FIGURE 8.—Organization of the information-management package.

For greatest efficiency, subroutines are organized into two classes: general and specific. General subroutines or "utilities" perform functions that may serve any segment of the monitor; three clusters of utilities are shown in figure 8. Terminal-utility subroutines provide for translations between computer-code structure and terminal-code structure during transmissions, establish communications-control sequences for polling and addressing, etc. The search-utility cluster includes all the subroutines that search the files for specific records or parts of records. This is a necessary function in order to retrieve from as well as update the file. The third cluster in the utility class is a miscellaneous group that performs a variety of generally useful data-manipulations.

Each of the specific subroutines, also illustrated as clusters, is tied to only one segment of the monitor. The purposes of the sort, add, change, and delete subroutines and the logic and action options are best understood as a function of the data-retrieval and update techniques described in the following sections.

The Data-Retrieval Technique

The Data-Retrieval Technique (DART) is an organized man-machine communication or "language" by which the user defines the specific information of interest and indicates what is to be done with those data. The communication may be conducted as an on-line "conversation," employing a terminal device connected to the computer, or as an indirect "batch" operation using punched cards or other preprepared media. Several versions of DART have been implemented, and two have been reported (refs. 1 and 2). The current version under development offers an expanded set of options, increasing control by the user.

The basic mechanism of DART is the question-answer technique similar to that used in STAT.

The computer, acting through the programmed management system (IMP), poses questions to the user. Answers entered in response to these questions provide the definitions and the action commands which control the programmed manipulation of the records on file. Therefore, two sets of responses are required from the user: (1) a definition of the data of interest, and (2) an instruction indicating the desired action to be performed on the defined data.

Definition is accomplished by the same set of questions that define each report in the file: SS NO, RECORD, TYPE, and DATE. For example, in answer to the question SS NO, the user may enter a particular social security number. RECORD is answered by the name of the record of interest, such as SURVEY. The response to TYPE is the name of one of the subdivisions of the survey system (table 1), and DATE is answered by entry of a specific day, month, and year. The user may enter the word ALL to indicate that a definition category is to be searched completely.

Figure 9 illustrates the retrieval of a complete VISION SURVEY. In this and all subsequent illustrations, underlining indicates the information entered by the user. Each response from the computer is initiated by a line of printing that re-

```
SS NO      123-45-6789.
RECORD     SURVEY.
TYPE       VISION.
DATE       ALL.
ACTION     LIST.

123-45-6789      SURVEY       VISION        01 JAN 68
VISUAL ACUITY
 DISTANT
  UNCORRECTED
   OD                    20/15
   OS                    20/15
 NEAR
  UNCORRECTED
   OD                    20/15
   OS                    20/17
ACCOMMODATION
 OD                      7.2
 OS                      7.3
INTRAOCULAR TENSION      15.3 MM HG O U
VISUAL FIELDS
 CONFRONTATION           NEG
DEPTH PERCEPTION
 TEST USED               VTA-ND
 UNCORRECTED             PASSES
HETEROPHORIA             DISTANCE    ES     EX     RH     LH
 1                       20'          0      0      0      0
 2                       13"          0      0      0      0
NPC
PD
PRISM DIV
PRISM CONV
COVER TEST               ORTHO
RED LENS TEST
NIGHT VISION             NIBH
COLOR VISION             ON RECORD AS PASSES
DIAGNOSES
COMMENTS
EXAMINER                 GEORGE L DAILY, MD
TYPED                    MM/01FEB68
```

FIGURE 9.—Vision survey retrieved.

states the definition data for the report retrieved.

The user may not want to retrieve a complete report but only a portion or even a single item. In order to limit the amount of data retrieved, the user may specify a portion or a single item of a report by adding the desired heading terms after the TYPE definition. For example, in answer to the question TYPE, the user may enter not only the name of a report (e.g., XRAY) but also an item heading (e.g., DIAGNOSIS):

TYPE XRAY—DIAGNOSIS

Now the search program will look within the body of the report and retrieve only the specified heading DIAGNOSIS and any associated data.

The user may further define the retrieval criteria by limiting the data of interest:

TYPE XRAY—DIAGNOSIS: NEG

In this instance retrieval occurs only if the report searched has the term NEG or NEGATIVE stored in the data field associated with the heading DIAGNOSIS.

This capability is designated "strong search" logic; it allows the user to specify any series (or string) of alphabetic or numeric characters to be used as a model during the computer search of the stored files. The user indicates whether a specific string is a heading or data by utilizing the colon (:) just as it is used in the basic file structure. A string preceded by a colon is understood to be data; the absence of a preceding colon indicates heading.

Boolean logic adds another dimension to the search definition by offering AND/OR/NOT control. Several sets of strings, headings with or without data, may now be linked by these terms into complex retrieval parameters. AND linkage requires that both conditions be met for constitution of a valid retrieval. OR linkage is satisfied if either condition is met. NOT allows invalidation based on a specified characteristic. For demonstration of the use of Boolean logic in MEDATA, assume the files to contain information on only two patients whose height is exactly 72 in. each. One of these and a third patient each weigh exactly 165 lb. Therefore only one of these three men is both 72 in. and 165 lb in height and weight. Figures 10 to 12 illustrate application of the AND/OR/NOT logic to such a file.

```
SS NO     ALL.
RECORD    SURVEY.
TYPE      MEASUREMENT-HEIGHT: 72 IN AND WEIGHT: 165 LB.
DATE      ALL.
ACTION    LIST.

123-45-6789   SURVEY   MEASUREMENT   01 JAN 68
GENERAL
 HEIGHT                 72 IN
 WEIGHT                 165 LB
```

FIGURE 10.—Example of *AND* logic.

```
SS NO     ALL.
RECORD    SURVEY.
TYPE      MEASUREMENT-HEIGHT: 72 IN OR WEIGHT: 165 LB.
DATE      ALL.
ACTION    LIST.

123-45-6789   SURVEY   MEASUREMENT   01 JAN 68
GENERAL
 HEIGHT                 72 IN
 WEIGHT                 165 LB

[redacted]    SURVEY   MEASUREMENT   16 NOV 67
 HEIGHT                 72 IN

[redacted]    SURVEY   MEASUREMENT   06 JAN 68
 WEIGHT                 165 LB
```

FIGURE 11.—Example of *OR* logic.

```
SS NO     ALL.
RECORD    SURVEY.
TYPE      MEASUREMENT-HEIGHT: 72 IN AND WEIGHT: NOT 165 LB.
DATE      ALL.
ACTION    LIST.

[redacted]    SURVEY   MEASUREMENT   16 NOV 67
 HEIGHT                 72 IN
```

FIGURE 12.—Example of *NOT* logic.

A third logic option—ranging—provides limited arithmetic latitude such that a range of values may be specified. A particular item satisfies the retrieval parameters if its numeric data value falls within the stated range. Figure 13 shows a simple request for a search of all individual LABORATORY SURVEY reports from any date to find those having white blood counts (WBC) between 4500 and 4700/mm^3. For this example only one case is retrieved. String, Boolean, and ranging logic may be mixed in a single definition set (fig. 14).

After the user has defined the specific information of interest, he indicates what the computer is to do with it. The ACTION commands available to the user begin with LIST; figures 9 to 14 have demonstrated the results of this command. LIST simply directs the computer system to produce the defined data in a standardized format.

DART takes maximum advantage of the facsimile storage of information. The basic format in which data are returned to the requestor is identical with the input format. Output is essentially a "dump" of the stored material as exemplified in figure 9. Whenever a definition specifies less than a complete report, as is usually the case, a modification called diagonal-retrieval formatting is employed. Diagonal-retrieval formatting refers to recovery of all superior and all inferior headings related to each defined heading. Figures 10 to 14 show that this format represents facsimile retrieval with all extraneous information removed. The purpose of diagonal formatting is to preserve the most complete meaning of retrieved information. Experience has verified the value of this policy even though sometimes unnecessary supplementary data are returned. An example of extra (but in this case probably useful) retrieval is the differential count data printed with the WBC in figure 13—because the original organization of headings for the laboratory-survey form placed the differential count one hierarchical level under the white blood count.

```
SS NO     ALL.
RECORD    SURVEY.
TYPE      LABORATORY-WBC: 4500 TO 4700.
DATE      ALL.
ACTION    LIST.

123-45-6789   SURVEY   LABORATORY   01 JAN 68
HEMATOLOGY
 CBC
  WBC                  4600
   DIFF
    NEUT               56
    EO                 01
    BASO               00
    LYMPH              43
    MONO               00
```

FIGURE 13.—Example of ranging logic.

```
SS NO     ALL.
RECORD    SURVEY.
TYPE      LABORATORY-VDRL:NEG AND (HB: 15 TO 16 OR HCT: NOT 0 TO 45).
DATE      ALL.
ACTION    LIST.

123-45-6789   SURVEY   LABORATORY   01 JAN 68
IMMUNOLOGY
 VDRL                  NEG
HEMATOLOGY
 CBC
  HB                   15 GM

[redacted]    SURVEY   LABORATORY   16 NOV 67
 VDRL                  NEG
HEMATOLOGY
 CBC
  HB                   16 GM
  HCT                  47%

[redacted]    SURVEY   LABORATORY   06 JAN 68
 VDRL                  NEG
HEMATOLOGY
 CBC
  HCT                  49%
```

FIGURE 14.—Example of combination of string-search, Boolean, and ranging logic.

Besides LIST, new action options now being added include COUNT, which totals the number of valid records matching the retrieval parameters; SPECIAL PRINT FORMATS, which are a group of output programs providing for special-purpose formatting of the output of a retrieval request; and several miscellaneous options to perform other special tasks. Still another kind of action command, not illustrated, indicates the particular output instrument to be used, such as terminal, line printer, magnetic tape, paper tape, punched cards, or console typewriter.

The Update Techniques

The MEDATA System recognizes three levels of need for capability in updating. The basic update technique adds or deletes complete reports. A new report may be added to the established file by placing an "A" in a specific location in the leader portion of a data tape prepared by STAT methodology. Since definition of the report is already present in the data tape, recognition of the "A" allows the program to sort and file that report. Deletion of a complete report uses a similar approach: "D" is positioned in the leader of a data tape carrying a complete definition of that specific report; here, however, no body to the report is required.

The peripheral update technique is the second level of capability used to modify information already stored in the file; examples include correction of errors or addition of new data. This level is designed for use by the original preparers of the data tapes for computer input and for users of the data-retrieval technique. The peripheral update technique allows modifications only in the data fields of the file. This restriction is necessary because the format and organization of all reports in a file must remain consistent; otherwise information could be "lost" within the file by changing of the standard headings of their relations.

In operation the peripheral update technique is parallel with DART. In the MEDATA system there are two reasons for recalling material stored in the file. The first and most frequent reason is retrieval of information; the mechanism has been discussed in detail. The second reason is updating by addition to, change in, or deletion of data within the file. To accomplish these actions one must first define the specific information of interest. The peripheral update, therefore, uses the same communications scheme as does DART, with change in only the action-command options. The update-action commands are: ADD, followed by the new information to be added; CHANGE, followed by new information which is to replace the current contents of the defined data field; and DELETE, which needs no following character string but automatically erases the present contents of the defined data field.

The third level or central update technique remains in the development stage. Its function is accommodation of generalized changes in file structure to allow for uniform reorganization of a report format within a given file, to perform a universal alteration on one or more headings throughout the file, or to permit any other necessary modification of the file that is not provided for by the basic and peripheral techniques. Because of its broad power, the central update capability should be controlled by those responsible for the filing system; it should not be available directly to the general user.

THE FUTURE OF MEDATA

The MEDATA System is an operational, computerized, medical-record system, one version of which was to be installed at Manned Spacecraft Center by June 30, 1968. Additional development will provide general and specific programs for data analysis. Programming for terminal communications depends on the availability of computer equipment capable of multiterminal time-sharing. The most significant future contribution, however, will be ability to provide efficient support for continuing improvement of basic medical documentation.

REFERENCES

1. Farley, W. E.; Pulido, A. H.; Minckler, T. M.; and Cady, L. D., Jr.: Man-Machine Communications in the Biological-Medical Research Environment. Proceedings of 21st National Conference of the Association for Computing Machinery, 1966, pp. 263–267.
2. Horton, C. M.; Minckler, T. M.; and Cady, L. D., Jr.: MEDATA: A New Concept in Medical Records Management. Proceedings of Fall Joint Computer Conference, 1967, pp. 485–489.

Page intentionally left blank

CHAPTER 6

DEVELOPMENT OF A COMPUTER PROGRAM FOR AUTOMATED RECOVERY OF LABORATORY DATA

Edward C. Knoblock, Mervyn R. Stein, and Stephen T. Paine

The analysis of data, including data compilations, arithmetic and statistical computations, graphical presentations, and data tabulations, is an integral part of laboratory operation. A computer system capable of efficient storage and uncomplicated recovery, with the facility for logical and arithmetic manipulations, would be of great benefit to the laboratory. In a space-medicine program, where laboratory data are to be analyzed, the number of critical variables increases so significantly that handling and analysis of data become an even more complicated task.

Medically Oriented Language (MEDOL), an information system designed around laboratory data, was developed to accomplish these objectives. The data from the Project Mercury program served as a framework for development of a laboratory-oriented file structure for the MEDOL system, with emphasis on current space programs. This system includes flexibility for addition of new classes of data, and provides for storage and retrieval of data without requiring the user to be familiar with sophisticated computer languages. All critical parameters required for evaluation of quality control and to describe essential detail of each procedure are described within the file structure.

The MEDOL system employs a higher-level interpretive source language that requires the use of a few English-language-oriented statements for file design and entry and retrieval of data. A synonym capability is available through a system glossary, and a dimensional library provides for interconversion of units for the convenience of the user.

There are two modes of operation in the system: file-generation additions and deletions of data and format changes are handled under the Maintenance Mode; the Query Mode provides for retrieval, utilizing arithmetic and logical operations, and the generation of reports in tabular or free-form output. Provision was made for the protection of proprietary data by prevention of inadvertent release by unauthorized sources.

MEDOL is a third-generation information-processing system that originated from SISTRAN (System Information Storage Retrieval and Analysis). SISTRAN was specifically designed as a library system for storing, retrieving, and analyzing aerospace documents; it provides the basic system programs for MEDOL. Enhancements have been provided for the specific requirements for processing of biomedical data.

The MEDOL System is written primarily in FORTRAN IV, with a few macroassembly program (MAP) subroutines. It operates on a 7094 tape system or a 7094/7040 direct-couple system under IBSYS; a 16-tape drive system is recommended for maximum efficiency. MEDOL was designed to be independent of secondary storage. In addition, it is completely modular and open-ended to facilitate conversion to other computer configurations including on-line computer systems; it was implemented in an English-like syntactical and semantic source language to allow the user to communicate with the system with maximum ease.

The basic system provides functions or procedures for (1) generating files, updating files, and retrieving data from these files; (2) generating libraries and glossaries; (3) performing elementary arithmetic operations on data; (4) extracting information for array grouping; and (5) output of data.

SOURCE LANGUAGE

An information-processing system must be accessible to many users, each having different requirements, a significant number of whom have little computer experience. Therefore, the system must provide a language (source) that can be successfully handled by the inexperienced as well as the experienced. The advantage of employing an English-like source language is that the tasks to be performed by the computer can be declared by individuals who are closest to the data, even though their computer experience is limited.

The MEDOL source language was designed to meet this objective; it has many similarities to FORTRAN but is certainly easier to learn and use. Furthermore an individual familiar with FORTRAN will learn MEDOL faster, but this is not an absolute requirement. The language consists of key words and symbols combined into statements that describe the MEDOL procedures. The key words and statements are readily comprehended by the user since they are terms in common usage. Some representative procedures employed in the MEDOL source language are ERASE, ADD, DELETE, PRINT, and END OF DATA. The procedures are combined to form the source-language program (the mechanism by which the user indicates to the computer what jobs are to be performed). The program layout is intended to approximate the logical thought processes normally employed by the scientific investigator. Finally the user is relieved of many bookkeeping chores imposed by many other languages (e.g., FORTRAN).

Data entry to the MEDOL system is via the source-language program; they are entered as a continuous string. The items of data are separated by delimiters and arranged according to the file's tree structure; absent data are indicated by embedded commas. A hazard of this scheme is that omission of a comma can result in incorrect entry of data for all items following the error.

PREPROCESSOR AND PROCESSOR

One of the most significant features of the MEDOL system is the modularity, which is most apparent in the operation of the preprocessor and processor. Once the source program is read into the computer, the preprocessor analyzes the statements, breaks them into components, and finally codes them in the internal language as procedural directives; this is accomplished before and separately from execution of the program. If an error is detected in a statement, an appropriate error message accompanies the source-language listing, and a procedural directive is not generated.

The procedural directives supply the processor with information as to which programs to call. As part of the program-execution the processor must relate operations, locations, and references from one part of the program to those in another. This latter activity contrasts with the work of incremental compilers used for multiprocessing and on-line systems, where a statement is compiled and machine-language instructions are generated independently of the next statement.

The generation involves two aspects: format design and data input. The file is based on a tree-structure form. The user defines all his variables and assigns appropriate hierarchical relations between them for creation of the data file tree. The tree can contain up to 15 levels with a variable number of attributes within each. The data-bearing elements are at the lowest level of each branch.

In the "update" function, files, data in files, and formats of files can be altered. Format can be modified to correspond with a change in the data status. Attribute names may be added as new data are available and may be deleted when no longer needed. Data rearrangements within data strings may also be accommodated. Procedural directives generated in internal language by the preprocessor are executed by the processor.

MODES OF OPERATION

The system operates in two modes, the data-maintenance mode and the query mode, only one of which can be processed at any one time. Some of the procedures used are unique to each mode, while others may be used in common. The maintenance mode provides an input function for entry of source-language information and data whereby files are set up and updated, and dictionaries, glossaries, and libraries are provided in the data bank.

The query mode is used to obtain information from the data bank for the purpose of computation or output of data as a printed report, punched cards, or magnetic tape—retrieval may be by item, group, or file. It is initiated by referencing of data name(s) to the levels of structure desired. Three basic types of queries exist: nonarithmetic, arithmetic, and logical. The nonarithmetic operation functions for retrieval of data from either the temporary or permanent files.

The arithmetic part performs computations by use of the five basic arithmetic operators: exponentiation, multiplication, division, addition, and subtraction. In addition, FORTRAN IV functions can be utilized to augment the arithmetic capability. Special subroutines are used for statistical analysis.

The logic query provides for decision-making capability such as testing of the data bank against certain conditions. These expressions can be nested up to 15 levels to provide complex conditional expressions. In addition, arithmetic and nonarithmetic statements can be used within conditional expressions.

MULTIPLE FILES

The MEDOL system is capable of dynamic array. Data can be stored and retrieved in multidimensional arrays located in a temporary storage area. Single elements, rows, or entire arrays can be retrieved with one reference. The indexing scheme to identify individual elements and rows is simple to use, and index arithmetic is available. Arithmetic and logical computations can be performed on the data in the "hold" queue as well as in the permanent files. The hold queue is capable of unlimited storage since, when the core area is exhausted, data can be transferred to scratch tape.

REPORTS

Data can be delivered on punched cards, magnetic tape, or printed reports. The system provides a mechanism for "formatting" of the data for the three output media. Printed reports can be generated in free or formatted form (tabular form).

The file format was designed to accommodate aerospace laboratory and other biomedical data. The file design was based on the Mercury Program data, but the system is so flexible that subsequent data from the Gemini and Apollo programs can be easily added.

Although the system handles vital-signs, fluid-input, and food-input data, the file structure is primarily geared toward the processing of laboratory data. Factors common to the processing of routine laboratory data had to be considered in design of the file, as well as those unique to aerospace-data processing. Information concerning collection, storage, and transportation of specimens, including distribution to two or more laboratories, has to be included in an aerospace system. Each specimen must also be linked by data to the aerospace activities or experiments under way when the specimen was obtained.

The system contains three files as outlined in Appendix: the Astronaut History File, the Flight Data File, and the Test Description File. The History File lists the flight(s) in which the astronaut participated, his birth date, and comments; it can be expanded to include additional items of pertinence.

The Flight Data File contains the bulk of the biomedical data and includes the vital-signs, fluid-input, food-input, laboratory, and collection and storage data. In addition, the reference data fill a significant position in this file. The file is astronaut-oriented within flight.

The reference data cover the significant aerospace experiments associated with each flight and include a detailed chronological history of the flight. The data are divided into four basic periods: base-line, preflight, flight, and postflight. The base-line period begins when the astronaut enters the space program (for his first flight) or on completion of his last postflight period (for subsequent flights). The base-line period for the Mercury program began with the medical interrogation at Lovelace Clinic in 1959; it extended to 30 days before launch. The preflight, flight, and postflight periods last from −30 days to launch, from launch to splashdown, and from splashdown to +30 days, respectively. Within each period are the specific reference events that can be related to the laboratory specimens or other data. Access to the data may be via the basic periods or through specific events. The basic structure of the reference data operated equally well for the Gemini and Apollo data after only a few minor adjustments.

The Test Description File was assembled to provide the parameters that would be required in evaluation of the data collected. The purpose of this file was to provide a sufficiently comprehensive description of the entire laboratory procedure for comparisons between laboratories, and to determine whether observed responses were significant. For this purpose it was necessary to describe each type of sample, the analytical technique, the laboratory, the method, and collection and storage of the sample, and to identify performing personnel. Also included for method-description are the literature reference, the precision to be expected of the method, and a "normal" range with a gate to identify unusual results immediately. With the various tests then assembled by groups of associated analyses, the file allows ready access by the user to recover desired information. The subroutine of table of synonyms provides access without the user having to know the precise description of the test within the file.

During the process for data-retrieval the user describes the desired data with instructions for computer operations which follow the logical sequence of the input of data into the program. Table 1 shows the format of a representative query. By application of such approaches, the three files provided enable extensive evaluation of the experience under examination.

SUMMARY

The MEDOL system is not foolproof, nor does it think for the user; it is a tool for handling large numbers of medical data. If the user expends enough thought on how he wants to query the system and how he wants the medical data presented for examination, and if he devotes time to planning and laying-out of the files, tree struc-

TABLE 1.—*Format of a Representative Query*†

1. Extract enzyme (specify tests) test results
   a. For pre- and postflight periods.
   b. For astronaut × (ensure that the collection dates for these tests for this period for the astronaut are in absolute form).
2. Extract normal range for these tests.
3. Test for out-of-range values.
4. Print out all extracted results in the following format:

Astronaut: James E. Doe
Period: Pre- and postflight

Period	Date, 1962	Test			
		SGOT, I.U.	SGPT, I.U.	LDH, I.U.	ALK P., B.U.
Preflight	2 Jan	19	6	190	—
Preflight	8 Feb	27	10	125	—
Postflight	20 Apr	*68	*50	*560	11

or

Astronaut: James E. Doe
Period: Pre- and postflight

Test	Preflight dates, 1962		Postflight, 20 Apr 1962
	2 Jan	8 Feb	
SGOT, I.U.	19	27	*68
SGPT, I.U.	6	10	*50
LDH, I.U.	190	125	*560
ALK P., B.U.	—	—	11

† All numbers are integers without signs.

tures, and data formats, clear and useful reports can be obtained continuously.

During recent development of MEDOL, many difficult and unique problems had to be resolved. The system's information-retrieval and analytic structures already existed, but enhancements of this basic program were needed to accommodate the greater character strings typical of medical queries and files, to make updating operations easier, and to gain greater computational ability through additional subroutines.

Around this framework it was necessary to build a medically oriented language that would allow the investigator to file laboratory information and query the files in a language familiar to him. For this purpose, glossaries of medical synonyms and units-conversion tables were constructed and placed in the system's reference files. In addition, diagnostic messages designed to pinpoint errors in the data, queries, and file-entry operations were developed.

Finally the most difficult task was the setting up of file formats that would provide flexibility, beyond the Mercury-data base, for the Gemini and Apollo laboratory analyses. Factors involving multiple astronauts on multiple flights had to be taken into account. Since all evaluations were not completed in a single laboratory, parallel studies from two or more laboratories had to be suitable for separate identification and analyses.

The files had to be so structured that they could be maintained and searched efficiently. Therefore, they were structured in trees of hierarchies with nested "repeats" to accommodate multiple flights, laboratories, specimens, and comments.

The present system will not satisfy all future requirements within the time and cost constraints provided, but a basic operational system has been constructed. The next step should be provision of greater repeat-generating capability through plot subroutines and flexible data-association in tabular form, so that the medical examiner can scan the printout (i.e., graphic plots and columnar listing of associative attributes) and reach a correct conclusion quickly.

As more data become available, greater analytic capability must be built into the MEDOL system to provide statistical correlations between attributes of interest to a medical investigator. Finally the internal system should be made more machine-independent by elimination of the few assembly-language terms in the preprocessor and processor, which would provide for easy transfer of MEDOL from one computer system to another.

Page intentionally left blank

APPENDIX

NASA'S LABORATORY-DATA SYSTEM'S TREE STRUCTURE

I. Astronaut's History File:
- a) Name
- b) Serial number
- c) Flight number *implicit repeat
- d) Birth date
- e) Comment *implicit repeat

II. Flight-Data File:
- a) Flight number
- b) Astronaut
  - 1. Astronaut name
  - 2. Serial number
- c) Reference data
  - 1. Base-line period (from first data point, or end of last postflight period, to beginning of next preflight period)
    - a. Date/time
      - 1. Begin
      - 2. End
    - b. Lovelace Clinic
      - 1. Date/time
        - a. Begin
        - b. End
      - 2. Comment
    - c. Simulations *implicit repeat
      - 1. Date/time
        - a. Begin
        - b. End
      - 2. Presimulation period
        - a. Begin
        - b. End
      - 3. Simulation period
        - a. Begin
        - b. End
      - 4. Postsimulation period
        - a. Begin
        - b. End
      - 5. Name
      - 6. Comment
    - d. Diurnal studies
      - 1. Date/time
        - a. Begin
        - b. End
      - 2. Comment
  - 2. Preflight period (from −30 days to launch)
    - a. Date/time
      - 1. Begin
      - 2. End
    - b. Simulations *implicit repeat
      - 1. Date/time
        - a. Begin
        - b. End
      - 2. Name
      - 3. Comment
    - c. Preflight physical exam (not to include the exam during countdown) *implicit repeat
      - 1. Date/time
      - 2. Comment
    - d. Countdown, abort *implicit repeat
      - 1. Date/time
        - a. Begin
        - b. End
      - 2. Awaken
        - a. Date/time
        - b. Comment
      - 3. Prelaunch breakfast
        - a. Date/time
        - b. Comment
      - 4. Don pressure suit
        - a. Date/time
        - b. Comment
      - 5. Insertion into spacecraft
        - a. Date/time
        - b. Comment
      - 6. Flight cancelled
        - a. Date/time
        - b. Comment
    - e. Countdown, flight (from awaken to launch)

1. Date/time
   a. Begin
   b. End
2. Awaken
   a. Date/time
   b. Comment
3. Prelaunch breakfast
   a. Date/time
   b. Comment
4. Don pressure suit
   a. Date/time
   b. Comment
5. Insertion into spacecraft
   a. Date/time
   b. Comment

3. Flight period (from launch to splashdown)
   a. Date/time
      1. Begin
      2. End
   b. Lift-off (from ignition to insertion into orbit)
      1. Date/time
         a. Begin
         b. End
      2. Comment
   c. In-flight (from insertion to firing of retro-rockets)
      1. Date/time
         a. Begin
         b. End
      2. Comment
   d. Reentry (from retro-rockets to splashdown)
      1. Date/time
         a. Begin
         b. End
      2. Comment
4. Postflight period (from splashdown to +30 days)
   a. Date/time
      1. Begin
      2. End
   b. Recovery (from splashdown to debriefing site)
      1. Date/time
         a. Begin
         b. End
      2. Recovery site (Atlantic or Pacific Ocean)
      3. Comment *implicit repeat
   c. Debriefing (from arrival to release)
      1. Date/time
         a. Begin
         b. End
      2. Period *implicit repeat
         a. Date/time
            1. Begin
            2. End
         b. Place (USA, aircraft carrier, Grand Turk Island, or Grand Bahama Island)
         c. Comment
   d. Additional activities *implicit repeat
      1. Date/time
         a. Begin
         b. End
      2. Name
      3. Comment

(Under Base-line, Preflight, Flight, and Postflight periods are listed some of the reference points for the Mercury program. The list is by no means complete and will continue to be enlarged for the Gemini and Apollo programs.)

d) Vital signs
   1. Temperature *implicit repeat
      a. Date/time
      b. Value (°F)
      c. Anatomic site (oral, rectal, or axillary)
      d. Comment
   2. Blood pressure *implicit repeat
      a. Date/time
      b. B.P. data *implicit repeat
         1. Value (mm-Hg)
         2. Arm (right or left)
         3. Position (supine, standing, sitting)
         4. Comment
   3. Weight *implicit repeat
      a. Date/time
      b. Value (lb)
      c. Status (after voiding and/or nude) *implicit repeat
      d. Comment
   4. Pulse *implicit repeat
      a. Date/time

- b. Pulse data *implicit repeat
  1. Position (supine, standing, sitting)
  2. Value (beats per minute)
  3. Exercise status (before or after exercise)
  4. Comment
5. Respiration *implicit repeat
   - a. Date/time
   - b. Value (breaths per minute)
   - c. Comment
6. Extremity measurement *implicit repeat
   - a. Date/time
   - b. Extremity data *implicit repeat
     1. Value (in.)
     2. Extremity site (wrist, forearm, thigh, calf, and ankle)
     3. Side (left or right)
     4. Comment
7. Vital capacity *implicit repeat
   - a. Date/time
   - b. Value (liters)
   - c. Comment
8. General comments *implicit repeat
   - a. Date/time
   - b. Comment

e) Fluid input
1. Fluid data *implicit repeat
   - a. Date/time
     1. Begin
     2. End
   - b. Volume (ml)
   - c. Type (water, tea, suspended food, soup, coffee, juice, other, or combinations) *implicit repeat
   - d. Comment

f) Food input
1. Food data *implicit repeat
   - a. Date/time
     1. Begin
     2. End
   - b. Type (e.g., apples, potatoes) *implicit repeat
   - c. Elemental ingredients (e.g., vitamins, calcium) *implicit repeat
   - d. Meal
   - e. Comment

(Food data for Mercury program should be entered under "Meal" and "Type"; "Elemental ingredients" will be used in Gemini and Apollo programs, but not for Mercury.)

g) Specimen collection and storage *implicit repeat
1. Specimen (e.g., blood)
2. CS data *implicit repeat
   - a. Collection interval
     1. Begin
     2. End
   - b. Collection (refers to the entire sample) *implicit repeat
     1. Sample collection date/time (end time for urine)
     2. Centrifugation date/time (for blood specimens)
     3. Volume (for urine specimens)
     4. Personnel
     5. Comment
   - c. Distribution (a description of division of the samples and distribution to their performing laboratories) *implicit repeat
     1. Sample [the types of blood samples that were sent to the lab. (2, below) from this collection; serum, plasma, or whole blood] *implicit repeat
     2. Laboratory (performing)
     3. Storage *implicit repeat
        - a. Phase (initial and final, or blank)
        - b. Storage method (deep freeze $-10°$ F, deep freeze $-40°$ F, dry ice, or liquid N_2)
        - c. Date/time
          1. Begin
          2. End
        - d. Comment (e.g., sample thawed or lost)
     4. Transport (includes transportation of the sample from the collection-storage area to the distribution point, and from the latter to the performing laboratory; or from the collection-storage area directly to the performing laboratory) *implicit repeat

a. Phase (initial, final, or direct)
b. Storage method
c. Date/time
    1. Begin
    2. End
d. Distribution point (applies only to initial phase; WRAIR in this case)
e. Comment

h) Laboratory Test Performance
1. Blood
    a. Chemistry
        1. Electrolytes
            a. Sodium *implicit repeat
                1. Collection date
                2. Result
                3. Comment
                4. Laboratory (leave empty for Mercury program)
                5. Performance date
            b. Potassium *implicit repeat
                .
                .
                .
                .
        2. Enzymes
            .
            .
            .
            .
        3. Catecholamines
            .
            .
            .
            .
        4. Minerals
            a. Cations
                .
                .
                .
                .
            b. Anions
                .
                .
                .
                .
        5. Proteins
            a. Total protein *implicit repeat
                .
                .
                .
                .
            b. Electrophoresis *implicit repeat
                1. Collection date
                2. Fractions
                    a. Albumin
                    b. Alpha–1 globulin
                    c. Alpha–2 globulin
                    d. Beta–1 globulin
                    e. Beta–2 globulin
                    f. Gamma
                3. Comment
                4. Laboratory
                5. Performance date
        6. Steroids
            .
            .
            .
            .
        7. Blood gases
            .
            .
            .
            .
        8. Miscellaneous
            .
            .
            .
            .
    b. Hematology
        1. Routine *implicit repeat
            a. Collection date
            b. Tests
                1. Hemoglobin
                    a. Result
                    b. Comment
                2. Hematocrit
                    a. Result
                    b. Comment
                3. WBC

a. Result
b. Comment
4. RBC
a. Result
b. Comment
5. WBC differential
a. Lymphocytes
b. Neutrophils
c. Stabs
d. Monocytes
e. Eosinophils
f. Basophils
g. Comment
6. RBC morphology (descriptive)
7. Platelets
a. Result
b. Comment
c. Laboratory
d. Performance date
2. Miscellaneous
a. ESR *implicit repeat
. . . .
c. Serology
1. VDRL *implicit repeat
. . . .

[All blood tests not described with a special format are to be treated with the standard format (see Sodium, above).]

2. Urine
a. Urinalysis *implicit repeat
1. Collection date
a. Begin
b. End
2. Tests
a. Volume
b. Specific gravity
c. pH
d. Albumin
e. Glucose
f. Ketones
g. Occult blood
h. Bile
i. Microscopic

3. Comment
4. Laboratory
5. Performance date
b. Electrolytes
1. Sodium *implicit repeat
a. Collection date
1. Begin
2. End
b. Result
c. Comment
d. Laboratory
e. Performance date
. . . .
c. Catecholamines
. . . .
d. Minerals
1. Cations
. . . .
2. Anions
. . . .
c. Steroids
. . . .
f. Diurnal studies *implicit repeat
1. Collection date
a. Begin
b. End
2. Sample *implicit repeat
a. Collection period
1. Begin
2. End
b. Catecholamines
1. Test *implicit repeat
a. Name
b. Result

c. Comment
d. Performance date
c. Steroids
1. Test *implicit repeat
a. Name
b. Result
c. Comment
d. Performance date
g. Miscellaneous
1. Albumin *implicit repeat
.
.
.
.
2. Xylose absorption test *implicit repeat
a. Collection date
b. Condition (a, b, or c)
c. Result
1. H 1
2. H 2
3. H 3
4. H 4
5. H 5
d. Comment
e. Laboratory
f. Performance date

[All urine tests not described with a special format are to be treated with the standard format (see Sodium, above). The following tests have special formats:]

3. Bone marrow *implicit repeat
a. Collection date
b. Result
1. Metamyelocyte
2. Myelocyte–C
3. Myelocyte–B
4. Myelocyte–A
5. Myeloblast
6. Late erythroblast
7. Normoblast
8. Eosinophilic myelocyte
9. Plasma cells
10. Megakaryocytes
c. Comment
d. Laboratory
e. Performance date
4. Gastric analysis *implicit repeat
a. Collection date
b. Test meal
c. After ingestion (minutes)
d. Tests
1. Volume
2. Total acid
3. Free acid
4. Appearance (text)
e. Comment
f. Laboratory
g. Performance date
5. Stool *implicit repeat
a. Collection date
b. Results
1. Character
2. Direct
3. Concentrated
a. Faust
b. De Rivas
c. Comment
d. Laboratory
e. Performance date
6. Semen *implicit repeat
a. Collection date
b. Results
1. Volume
2. Sperm count
3. Motility
a. Motile
b. Moribund
c. Inert
4. Morphology
a. Abnormal
b. Normal
5. WBC
c. Comment
d. Laboratory
e. Performance date

III. Test Description File
1) Blood
a) Chemistry
1. Electrolytes
a. Sodium
1. Laboratory *implicit repeat
a. Name
b. Period
1. Begin
2. End

- c. Method
  - 1. Name
  - 2. Reference
  - 3. Precision (S.D. or %)
  - 4. Normal range
    - a. High
    - b. Low
  - 5. Sample
    - a. Type (serum, etc.)
    - b. Anticoagulant
  - 6. Preservative
  - 7. Comment
- d. Personnel *implicit repeat

b. Potassium

.
.
.
.

2. Enzymes

.
.
.
.

3. Catecholamines

.
.
.
.

4. Minerals
   - a. Cations

     .
     .
     .
     .

   - b. Anions

     .
     .
     .
     .

5. Proteins
   - a. Total protein

     .
     .
     .
     .

   - b. Electrophoresis
     - 1. Laboratory *implicit repeat
       - a. Name
       - b. Period
         - 1. Begin
         - 2. End
       - c. Method
         - 1. Name
         - 2. Reference
         - 3. Precision
         - 4. Normal range
           - a. Albumin
             - 1. High
             - 2. Low
           - b. Alpha–1
             - 1. High
             - 2. Low
           - c. Alpha–2
             - 1. High
             - 2. Low
           - d. Beta–1
             - 1. High
             - 2. Low
           - e. Beta–2
             - 1. High
             - 2. Low
           - f. Gamma
             - 1. High
             - 2. Low
         - 5. Sample
           - a. Type
           - b. Anticoagulant
         - 6. Preservative
         - 7. Comment
       - d. Personnel *implicit repeat

6. Steroids

.
.
.
.

7. Blood gases

.
.
.
.

8. Miscellaneous

.
.
.

b) Hematology
 1. Routine
  a. Laboratory *implicit repeat
   1. Name
   2. Period
    a. Begin
    b. End
   3. Hemoglobin
    a. Method
     1. Name
     2. Reference
     3. Precision
     4. Normal range
      a. High
      b. Low
     5. Comment
   4. Hematocrit
   5. WBC
   6. RBC
   7. WBC differential
    a. Method
     1. Name
     2. Reference
     3. Normal range
      a. Lymphocytes
       1. High
       2. Low
      b. Neutrophils
       1. High
       2. Low
      c. Stabs
       1. High
       2. Low
      d. Monocytes
       1. High
       2. Low
      e. Eosinophils
       1. High
       2. Low
      f. Easophils
       1. High
       2. Low
     4. Comment

(RBC morphology not included)
   8. Platelets
.
.
.
   9. Sample
    a. Type
    b. Anticoagulant
   10. Preservative
   11. Personnel *implicit repeat

(For HCT, WBC, RBC, and platelets, use the format under Hemoglobin.)
 2. Miscellaneous
  a. ESR
.
.
.
.

c) Serology
 1. VDRL
.
.
.
.

[All blood tests that do not have a special format are to be treated with the standard format (see Sodium, above).]

2) Urine
 a) Urinalysis
  1. Laboratory *implicit repeat
   a. Name
   b. Period
    1. Begin
    2. End
   c. Specific gravity
    1. Method
     a. Name
     b. Reference
     c. Precision
     d. Comment
   d. pH
.
.
.
   e. Microscopic
.
.
.
   f. Preservative

            g. Personnel *implicit repeat
    b) Electrolytes
        1. Sodium
            a. Laboratory *implicit repeat
                1. Name
                2. Period
                    a. Begin
                    b. End
                3. Method
                    a. Name
                    b. Reference
                    c. Precision
                    d. Normal range
                        1. High
                        2. Low
                    e. Preservative
                    f. Comment
                4. Personnel *implicit repeat
        2. Potassium
            .
            .
            .
    c) Catecholamines
        .
        .
        .
        .
    d) Minerals
        .
        .
        .
    e) Steroids
        .
        .
        .
    f) Miscellaneous (volume transferred to collection group)
        1. Albumin
            .
            .
            .
        2. Xylose absorption test
            a. Laboratory *implicit repeat
                1. Name
                2. Period
                    a. Begin
                    b. End
                3. Method
                    a. Name
                    b. Reference
                    c. Precision
                    d. Preservative
                    e. Comment
                4. Test condition *implicit repeat
                    a. Name (plus text)
                    b. Control data *implicit repeat
                        1. Control individual
                        2. H 1
                        3. H 2
                        4. H 3
                        5. H 4
                        6. H 5
                5. Personnel *implicit repeat
3) Bone marrow
    a) Laboratory *implicit repeat
        1. Name
        2. Period
            a. Begin
            b. End
        3. Method
            a. Name
            b. Reference
            c. Normal range
                1. Metamyelocyte
                    a. High
                    b. Low
                2. Myelocyte-C
                    a. High
                    b. Low
                3. Myelocyte-B
                    a. High
                    b. Low
                4. Myelocyte-A
                    a. High
                    b. Low
                5. Myeloblast
                    a. High
                    b. Low
                6. Late erythroblast
                    a. High

b. Low
7. Normoblast
a. High
b. Low
8. Eosinophilic myelocyte
a. High
b. Low
9. Plasma cell
a. High
b. Low
10. Megakaryocyte
a. High
b. Low
d. Comment
4. Personnal *implicit repeat
4) Gastric analysis
a) Laboratory *implicit repeat
1. Name
2. Period
a. Begin
b. End
3. Test meal
4. After ingestion
5. Total acidity
a. Method
1. Name
2. Reference
3. Precision
4. Normal range
a. High
b. Low
5. Comment
6. Free acidity
a. Method
1. Name
2. Reference
3. Precision
4. Normal range
a. High
b. Low
5. Comment
7. Personnel *implicit repeat
5) Stool
a) Laboratory *implicit repeat
1. Name
2. Period
a. Begin
b. End
3. Direct
a. Method
1. Name
2. Reference
3. Comment
4. Concentrated Faust
a. Method
1. Name
2. Reference
3. Comment
5. Concentrated De Rivas
a. Method
1. Name
2. Reference
3. Comment
6. Personnel *implicit repeat
6) Semen
a) Laboratory *implicit repeat
1. Name
2. Period
a. Begin
b. End
3. Method
a. Name
b. Reference
c. Normal range
1. Sperm count
a. High
b. Low
2. Motility
a. Motile
1. High
2. Low
b. Moribund
1. High
2. Low
c. Inert
1. High
2. Low
3. Morphology
a. Abnormal forms
1. High
2. Low
b. Normal forms
1. High
2. Low
4. WBC (occasional or absent)
d. Comment
4. Personnel *implicit repeat

CHAPTER 7

CONTINUOUS MONITORING AND INTERPRETATION OF ELECTROCARDIOGRAMS FROM SPACE

Cesar A. Caceres, Anna Lea Weihrer, Sidney Abraham, David E. Winer, Larry K. Jackson, Jerome Sashin, Stuart W. Rosner, Juan B. Calatayud, James W. McAllister, and Howard M. Hochberg

Monitoring of subjects in real time—that is, in stressful situations such as during intensive care before or after surgery, or in space capsules—can be of life-preserving necessity. Its chief limitation is that the real time of the subject is the same as that of the human monitor—generally a highly trained physician or nurse whose time is too scarce and costly for routine, noncreative use.

These aspects of monitoring raise timely questions for medical engineering: "To what extent can a computer system facilitate the work and save the time of a human medical monitor?" "Can a computer system provide rapid and reliable analysis of continuous data to relieve the human monitor of routine, noncreative tasks?"

From limited but varied experimentation in the automated monitoring of medical signals—including such vital parameters as the electrocardiogram (ECG), the electroencephalogram (EEG), heart rates, heart sounds, and respiratory curves—we have evidence that an automated system can transmit continuous signals, convert them to computer input, analyze them rapidly and accurately, and display them as needed during care of a patient, for other clinical or experimental purposes, or during space flight.

ON-LINE ANALYSIS OF LIMITED-TIME SIGNALS

On-line analysis of a signal, recorded within a defined time period, preceded on-line analysis during open-ended time periods. As a means of demonstrating the feasibility of on-line processing of conventional time-limited medical signals by computer, the Medical Systems Development Laboratory (MSDL, formerly the Instrumentation Field Station) has conducted several on-line-analysis projects since 1961. In most of these the telephone system has been used for transmission of the wave forms as analog signals to the computer site; there they are received by dataphone and put through an analog-to-digital converter, and thence go to a digital-computer system for analysis. The signals are processed, interpreted, and available for retransmission to the physician within seconds of receipt at the processing center. The measurements and interpretation are delivered to a teletypewriter or remote printer for the display. Figure 1 is a facsimile of a routine printout; to demonstrate the routine function of an automated system for all medical signals, electrocardiographic signals have been used as a model.

A computer-processed ECG is intended to be a diagnostic aid to the physician interpreting an ECG for patient care or clinical investigation. As such the computer interpretation should be viewed as a screening tool providing consistent answers when defined wave patterns are present. A computer interpretation, like that of the electrocardiographer, carries a differing weight in a patient's diagnosis depending on the clinical circumstances. Only the physician in charge of the patient can determine the weight to be given to the interpretation.

The beginner studying electrocardiography will find that computer interpretations can be useful

MEDICAL SYSTEMS DEVELOPMENT LAB--HEART DISEASE CONTROL PROGRAM
COMPUTER PROCESSED ELECTROCARDIOGRAM
C ST CLINIC

NAME DIAGNOSIS
NUMBER TAPE 0309 OPTION 000 DATE 12-05-67 TIME 00--29--34
HEIGHT 75 WEIGHT 208 AGE 61 MALE MEDS UNKNOWN
BP HYPERTENSION SYSTOLIC 160 ABOVE OR DIASTOLIC 95 OR ABOVE

LEAD	I	II	III	AVR	AVL	AVF	V1	V2	V3	V4	V5	V6
PA	.06	.10	-.06	-.09	.07	.07	.04	.00	.08	.06	.04	.04
PD	.08	.08	.06	.10	.08	.08	.04	.00	.08	.05	.05	.04
P'A	.00	.00	.00	.00	.00	-.03	-.03	.00	.00	.00	.00	.00
P'D	.00	.00	.00	.00	.00	.02	.07	.00	.00	.00	.00	.00
QA	-.04	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
QD	.02	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
RA	1.01	.47	.10	.00	.91	.17	.05	.91	1.97	2.31	1.52	.95
RD	.06	.06	.03	.00	.11	.02	.02	.05	.06	.06	.07	.06
SA	.00	-.31	-.90	-.74	.00	-.52	-.39	-.21	-.22	-.15	-.09	.00
SD	.00	.05	.08	.05	.00	.05	.06	.02	.02	.02	.02	.00
R'A	.00	.00	.00	.09	.00	.00	.00	.00	.00	.00	.00	.00
R'D	.00	.00	.00	.03	.00	.00	.00	.00	.00	.00	.00	.00
ST	.12	.12	.12	.12	.12	.12	.12	.12	.12	.12	.12	.12
STO	-.01	-.04	-.03	.05	.00	-.05	.08	-.02	-.03	-.05	-.05	-.08
STM	-.04	-.01	.03	.03	-.02	.01	-.01	.04	.01	-.04	-.02	-.05
STE	-.02	.01	.03	-.01	-.01	.04	-.02	.04	.06	-.01	.01	-.04
TA	.21	.24	.06	-.23	.09	.16	-.22	.29	.42	.34	.24	.16
PR	.13	.15	.09	.17	.12	.18	.15	.00	.16	.13	.13	.13
QRS	.08	.11	.11	.08	.11	.07	.08	.07	.08	.08	.09	.06
QT	.41	.41	.33	.40	.40	.36	.41	.39	.41	.43	.42	.38
RATE	58	55	61	57	56	54	58	58	57	57	54	58
CODE	2	3	2	2	3	2	3	2	3	2	2	3
CAL	83	83	83	83	83	83	83	83	83	83	83	83

AXIS IN	P	QRS	T	Q	R	S	STO	ST-T	QRS-T
DEGREES	47	-20	37		-03	-49	256	141	57

1131 RATE UNDER 60 . BRADYCARDIA

8311 QRS AXIS RANGE -10 to -59 . LEFT AXIS DEVIATION

FIGURE 1.—Facsimile of a routine printout.

as a "self teaching" method as he reviews the computer interpretation of the ECG, refers to the list of criteria to determine the specific abnormalities on which each diagnosis is based, identifies the abnormal computer measurements, and inspects the ECG tracing to study the waves from which the measurements were made. Similarly the practitioner should first review the computer interpretation of the ECG and then inspect the tracing if necessary.

In our system the 12-lead ECG is sent by telephone or recorded on FM magnetic tape with a specially designed data-acquisition system. At the computer center the signals are received or played back from magnetic tape, sampled 500 times per second, digitized, and entered into computer memory. The duration of signals analyzed by the computer at one time is about 4 sec. The analysis has the following features:

Identification and measurements of waves—All waves of clinical significance are identified, and their amplitudes (A) are determined in millivolts and durations (D) in seconds. Wave terminology is in general conventional. It should be noted that the initial negative wave of a QRS-complex is termed a Q-wave when small and an S-wave when large. In the criteria used for infarcts, an S-wave preceded by an R of zero is equivalent to a Q-wave. The ST-segment duration and the ST-amplitude are measured at three points: the onset (STO), middle (STM), and end (STE).

If waves are absent or below an arbitrary "suppression level," zero values are printed for amplitude and duration. The following suppression levels are used: P = 0.025 mV—Q, 0.025 mV; T = 0.03 mV—R = 0.025 mV—S = 0.06 mV; T, T′, T = 0.03 mV—R, 0.1 mV; Rd, 0.05 sec—S, 0.017; Sd, 0.06 sec.

Axis determination of P, QRS, T, Q, R, S, STO, ST-T, and QRS-T—A pair of limb leads, either in both standard leads or in both augmented leads, is selected on the basis of amplitudes of the wave whose axis is being determined. Since amplitude-determination is more accurate from large waves, the pair of leads selected is such that the smaller of the two waves is as large as possible. For example, if the net QRS-amplitude in leads I and II is 0.70 and 0.20 mV, respectively, and in aVR and aVL, 0.36 and 0.34 mV, respectively, the latter pair is selected because it has the larger of the two smaller waves.

No axis is calculated unless both amplitudes equal or exceed the following values: P, ±0.05 mV; ST-onset, ±0.10 mV. The axis may be checked graphically by plotting the amplitudes on a hexaxial figure, drawing perpendiculars to the axis, and determining the intersection of the perpendiculars. The axis, clockwise from the positive horizontal through three quadrants (0 to 270 degrees) and the final quadrant, clockwise is −90 to 0 degrees.

Special information in the printout—In the code line on the printout, the number in each lead column is the number of the electrocardiographic complex analyzed (QRS, counting from the left) in 4 sec of recording. This information is useful if there is variation between heart cycles or if artifacts are present. Code letters may be printed also when difficulties in processing are encountered.

The recorded calibration pulse is considered by the computer program to be equivalent to 1 mV, and all amplitudes are corrected to this level. The calibration on the tracing is shown in the computer printout as a percentage. For example, a value of 105 means that the calibration on the recorder was 10.5 mm. An interpretation of under- or overstandardization is given when the calibration is more than 25 percent above or below 1 cm/mV.

Electrocardiographic interpretation—The interpretation is based on wave measurements, axes, and the other findings that are summarized in the code line, and on their interrelation. The interpretations are both descriptive and interpretive. In the left column are the wave abnormalities found in the ECG; in the right column is the interpretation based on these abnormalities.

The objectiveness of the system has marked advantages. First, human bias is eliminated. Second, since all output data are in digital form, they are immediately or subsequently available for display and clinical interrelation with other events. Third, methods for predictive statistical interpretation are established.

In one demonstration in 1965, 1500 conventional resting ECG's were sent from Las Vegas, Nevada, to Washington, D.C., and returned. In many cases the interpretations were available before the electrodes were removed from the subjects. In a continuing demonstration project, the outpatient clinic at Hartford Hospital, Hartford, Conn., daily sends tracings to Washington, D.C., for analysis and return of results. These and other demonstrations have laid the foundation for on-line, real-time computer analysis and monitoring of any medical signals.

The objectivity of the available methods suggested their use not only as the basis for screening techniques to detect disease, but also for monitoring subjects in intensive-care units, triggering alarm systems or control servomechanisms to start therapeutic measures if necessary, and evaluating subjects engaged in any activity such as exercise, training, or space flight. The methods being developed can be adapted to computers aboard spacecraft, to computers in communications satellites for international medical purposes, or to small "hospital size" computers. We present some samples of what is now being done and suggest what should be expected from combinations of techniques and displays.

DISPLAY METHODS

Continuous tabulation and verbal analysis—After further research, the conventional method of electrocardiographic analysis should be replaced by statistically significant electrocardiographic values interrelated with statistically significant values from other signals. But today's physicians best understand significant change in a monitored ECG by empiric diagnostic statements based on the conventional diagnostic tracing.

Our computer measures all the ECG components of the cardiac cycle that are necessary for these statements. The verbal interpretation is based on interrelation of the numerical values following the empiric basis used in medicine. The result is the same as if a cardiologist were personally interpreting the ECG's. The advantages of computer analysis are that the automated system works around the clock and provides precise measurements and printout interpretations at any desired time.

Table 1 lists verbal interpretations available in our current computer monitoring program. The exact criteria given are subject to change as experience dictates. New diagnoses can be added to meet the physician's needs.

In our initial on-line, real-time trials, the first instances in which complete electrocardiographic data from any human were monitored and interpreted on an on-line, real-time basis by a digital computer system, continuously monitored ECG's were sampled. Analog signals from an astronaut in Gemini 7 were obtained in real time and intermittently monitored at 30-to-60-min intervals over a 2-week period.

After that initial trial, the MSDL made immediate, automatic, computer analyses of the ECG's of astronauts in orbit in Gemini flights 8 to 12. In that project, intended to test for feasibility and to demonstrate the capability of on-line, real-time analysis, the MSDL received continuous data. The signals were first transmitted from the capsule to the nearest tracking station which telemetered them to Goddard Space Flight Center; thence they were transmitted by telephone as multiplexed analog signals to the computer center in Washington, D.C.

Four electrocardiographic leads, two from each astronaut, were available, but, because of the limited ability of the then-existing computer system, only two signals—one from each astronaut—were analyzed immediately; the other two were recorded on tape for postflight analysis.

TABLE 1.—*Electrocardiographic Diagnosis and Criteria Used to Establish Diagnosis*

P-Wave Abnormalities	
Tall P-waves; P-pulmonale	PA, ≥ 0.30 mV in any three time blocks
Short PR	PR, ≥ 0.11 sec; rate, ≤ 99 in four time blocks. Other two time blocks if present must be <0.13 sec
Prolonged PR-intervals; first-degree atrioventricular block	PR, 0.21 to 0.29 sec; PA, ≥ 0.05 mV in four time blocks
No atrial activity detected; rule out nodal rhythm, atrial fibrillation	All P-waves $=0$ in at least five time blocks analyzed
Conduction Defects	
QRS-prolongation, 0.13 sec; intraventricular conduction defect	QRS-duration, ≥ 0.13 sec in any four time blocks; no value <0.16 sec
QRS-prolongation; intraventricular conduction defect	QRS-duration, ≥ 0.14 sec in any three leads; no value <0.12 sec
ST-Segment Abnormalities	
Abbreviations	STO—amplitude of onset of ST-segment STM—amplitude of midpoint of segment STE—amplitude of end of ST-segment
Junctional ST-depression; rising ST-segment	STO, ≥ -0.10 mV negative; STM $>$STO. T, positive; ST segment, ≥ 0.08
Minor ST-depression; flat or falling ST	STO between -0.05 and -0.09 mV; STM$\geq$STO (negative) in any two time blocks. ST-segment, ≥ 0.08
Moderate ST-depression; flat or falling ST-segment	STO, ≥ -0.10 mV (negative); STM$\geq$STO (negative) in any two time blocks. ST, ≥ 0.08 any two time blocks
ST-elevation; rule out early repolarization	STO, ≥ 0.08 in any four time blocks; or STO, ≥ 0.06 in any five time blocks
Marked ST-displacement, 1+; possible current of injury	STO, ≥ 0.20 mV positive or negative; or STO, STM, STE, ≥ 0.12 mV positive or negative
Extreme ST-displacement	STO, ≥ 0.35 mV, positive or negative, in any two time blocks
T-Wave Abnormalities	
Negative T-waves, -0.10 to -0.49 mV	TA, -0.10 to -0.49 mV; QRS peak-to-peak amplitude, ≥ 0.51 mV; in at least three time blocks
Negative T-waves, -0.50 to -0.99 mV	TA, -0.50 to -0.99 mV; QRS peak-to-peak amplitude, ≥ 0.51 mV; in at least three time blocks. No values less negative than -0.25 mV
Negative T-waves, -1.0 to -1.49 mV	TA, -1.0 to -1.49 mV; QRS peak-to-peak amplitude, ≥ 0.51 mV; in at least three time blocks. No values less negative than -0.25 mV
Negative T-waves, -1.5 to -1.99 mV	TA, -1.5 to -1.99 mV; QRS peak-to-peak amplitude, ≥ 0.51 mV; in at least three time blocks. No values less negative than -0.50 mV
Negative T-waves, -2.0 to 2.49 mV	TA, -2.0 to -2.49 mV; QRS peak-to-peak amplitude, ≥ 0.51 mV; in at least three time blocks. No value less negative than -1.0 mV
Negative T-waves, -2.5 to -2.99 mV	TA, -2.5 to -2.99 mV; QRS peak-to-peak amplitude, ≥ 0.51 mV; in at least three time blocks. No value less negative than -1.5 mV

TABLE 1.—*Electrocardiographic Diagnosis and Criteria Used to Establish Diagnosis*—(Concluded)

T-Wave Abnormalities (Continued)	
Negative T-waves, ≥ -3.0 mV	TA, ≥ -3.0 mV (negative); QRS peak-to-peak amplitude, ≥ 0.51 mV; in at least three time blocks. No values less negative than -2.0 mV
QT-Abnormalities	
Prolonged QT	QT, from <0.43 sec; TA, ≥ 0.10 mV positive or negative. Rate, ≥ 61 in three time blocks
Short QT	QT, ≤ 0.29 sec. Rate, ≤ 99 in at least two time blocks; must be analyzed. If two blocks are less than 0.30 Dx not made
Voltage Abnormalities	
Low voltage	Peak-to-peak QRS-voltage, ≤ 0.50 mV in all of at least four time blocks analyzed
Defective data	QRS-duration of zero in all time blocks suppresses other diagnoses
Rhythms	
Ventricular rate, >100 in three or more time blocks: tachycardia	Heart rate, ≥ 101 in three or more time blocks
Ventricular rate, <60 in three or more time blocks: bradycardia	Heart rate, ≤ 59 in three or more time blocks
Variable RR-interval in four time blocks: rule out arrhythmia	"A" code appears in "code line" in any five time blocks; the block must not be low in voltage
Variable RR-interval in four time blocks; rule out artifact or atrial fibrillation	"A" code appears in "code line" in any five time blocks; no block can be low in voltage
Atypical QRS or artifacts in two or more time blocks: rule out premature contractions	"E" appears in "code line" in any two time blocks

Immediate eight-channel analysis should be achieved very soon.

Throughout Gemini 7, 215 separate samples were received. The eleven 3.7-sec samples shown in facsimile in figure 2 are typical of the teletyped (TWX) output after on-line, real-time processing.

Our computerized monitoring program initially considers only pattern-recognition aspects of signal analysis and provides numerical values. Where in the example in figure 2 significant values occur in the numerical values, a persistence of asterisks indicates those "abnormal" values. Asterisks identify values considered suspicious in normal clinical practice. In monitoring, these are important (i.e., not due to noise, movement, or artifact) only when sustained for defined time periods. Noise or artifacts occurring during a specific electrocardiographic complex can also be identified by similar means with adequate recognition programs.

Variable template—The first requirement of a monitoring program, different from a time-limited analysis, is establishment of criteria for change. To demonstrate what is meant, a facsimile of an output display from one of our developmental programs is shown in figure 3. At the onset of the recording in figure 3 the subject's heart rate was 123; at the end, it was 144. The measurements of PR, QRS, QT are followed by PA (P-amplitude), PD (P-duration), STO (ST-segment onset), and

TRANSACTION NR 178
GET 282-59-00
GMT 014-29-00
REV NR 177
GEMINI 7NA01
CARNARVON

INSTRUMENTATION FIELD STATION--HEART DISEASE CONTROL PROGRAM
COMPUTER MONITORED ELECTROCARDIOGRAM

GEMINI ASTRONAUT DATA

	1	2	3	4	5	6	7	8	9	10	11
PA	.00	.08	.09	.10*	-.04*	.09	.07	-.05*	.07	.07	.06
PD	.00	.10	.07	.09	.04	.09	.07	.09	.07	.08	.09
P'A	.00	.00	.00	.00	.06	.00	.00	.11	.00	.00	.00
P'D	.00	.00	.00	.00	.08	.00	.00	.09	.00	.00	.00
RA	.95	.95	.89	.90	.91	.93	.95	.92	.84	.92	.91
RD	.06	.06	.06	.06	.06	.06	.06	.06	.07	.06	.08
SA	-.43	-.40	-.44	-.38	-.39	-.43	-.42	-.42	-.38	-.37	-.39
SD	.05	.05	.05	.05	.04	.05	.04	.04	.04	.04	.04
ST	.07	.06	.08	.12	.07	.12	.07	.07	.05	.06	.12
STO	-.05	.04	.00	.01	.01	.04	.00	.00	.00	.01	.01
STM	-.02	.03	.02	.05	.04	.04	.01	.03	.01	.03	.07
STE	.01	.06	.06	.24	.05	.27	.04	.04	.01	.05	.24
TA	.32*	.39	.37	.37*	.37	.39*	.37	.40	.43	.39	.43
TD	.15	.16	.15		.15		.15	.15	.09	.18	
T'A	-.05	.00	.00	.00	.00	-.04	.00	.00	.00	.00	.00
T'D	.08	.00	.00	.00	.00	.06	.00	.00	.00	.00	.00
PR	,00	.12	.11*	.13	.20	.12	.12	.22*	.11*	.12	.13
QRS	.11	.11	.11	.11	.10	.11	.10	.10	.11	.10	.12
QT	.41	.34	.34	.33	.33	.40	.34	.33	.37	.35	.37
RATE	85	82	93	82	82	81	78	85	79	85	75

WITHIN NORMAL LIMITS

FIGURE 2.—Facsimile of 11 samples of teletyped output.

STM (ST-segment midpoint). "No significant change" refers to the amplitudes and durations, with the initial measurement set as a base line for comparison.

Thus with this program the computer can make statements about significant change from preceding beats including the presence of an arrhythmia, other electrocardiographic interpretations, noise, or artifacts. From these items, indications of trends to normality or abnormality can be given according to clinical terms or criteria developed for special needs. The output in figure 3 is not intended for graphic display in real clinical situations; it is shown only to demonstrate output possibilities. Portions of the display, such as heart-rate or arrhythmia designations, are suitable for physicians' monitoring needs in which some data-reduction technique must be introduced to relieve the responsible physician of interpretation of vast amounts of continuous data. Other portions, such as the repetitive "no significant change" statement, can be used to trigger alarms, lights, or other mechanisms. With the analyzed data of measurements of P, QRS, etc., subservient-computer statistical routines can be used for evaluation of current results or with previous data, or for analysis of portions of the data with greater scrutiny.

Continuous analysis of signals can easily produce a monumental pileup of data. The goal of data-reduction is to discard superfluous detail and to preserve meaningful information. A time-varying template of measurements is a simple, useful approach to solution of the problem. A

INSTRUMENTATION FIELD STATION--HEART DISEASE CONTROL PROGRAM
COMPUTER PROCESSED CONTINUOUS ELECTROCARDIOGRAM
GEMINI-TITAN 3

Measurements of initial time period

Rate	PR	QRS	QT	PA	PD	RA	RD	STO	STM	TA	TD
123	.10	.06	030	.05	.06	0.31	.06	.04	-0.02	0.06	.12

```
TIME     -26.3 SECS                  RATE 122   NO SIGNIFICANT CHANGE
TIME     -22.6 SECS                  RATE 120   NO SIGNIFICANT CHANGE
TIME     -18.9 SECS                  RATE 110   NO SIGNIFICANT CHANGE
TIME     -15.2 SECS                  RATE 119   NO SIGNIFICANT CHANGE
TIME     -11.5 SECS                  RATE 126   NO SIGNIFICANT CHANGE
TIME     -07.8 SECS                  RATE 132   NO SIGNIFICANT CHANGE
TIME     -04.1 SECS                  RATE ---
ARTIFACT RULE OUT ARRHYTHMIA
UNRECOGNIZABLE
TIME     03.3 SECS                   RATE 144   NO SIGNIFICANT CHANGE
TIME     07.0 SECS                   RATE 144   NO SIGNIFICANT CHANGE
TIME     10.7 SECS                   RATE 140   NO SIGNIFICANT CHANGE
TIME     14.4 SECS                   RATE 142   NO SIGNIFICANT CHANGE
TIME     18.1 SECS                   RATE 138   NO SIGNIFICANT CHANGE
TIME     21.8 SECS                   RATE ---
MISSING DATA RULE OUT ECTOPIC BEAT
TIME     25.5 SECS                   RATE 144   NO SIGNIFICANT CHANGE
```

FIGURE 3.—Facsimile of an output display.

template of measurements can serve as a standard to signal the occurrence of physiologically important change in wave form, but several templates should be available for use simultaneously as needed.

In figure 3 the measurements of the first beat served as the initial standard. If all subsequent measurements show no significant change, as in this case, only a statement to that effect appears on the printout after the measurements of the first beat. If one or more criteria for significant change are exceeded, the measurements of the "changed" beat can be displayed on the printout. These new measurements can serve as a new template of measurement values.

Safeguards are necessary to prevent introduction of a faulty "standard" template. Measurements other than the one(s) showing sustained significant change must be checked against the previous template of measurements for transient significant changes. In this way, short-lived alterations in measurement (physiologic or artifactual) do not affect the overall template but are still appropriately noted. For example, ventricular-rate variation can occur, suggesting artifact or arrhythmia, or ST-segment changes due to artifact, tachycardia, or myocardial injury. Whether the variation is transient or permanent, the cause is an important consideration before display to the monitoring physician.

Unfortunately there are no established guidelines—nor even consensus—regarding the significance of wave changes during rest or activity. In one of our efforts to develop tentative criteria, a group of 2200 electrocardiographic computer-measured recordings (12-lead ECG's) were analyzed. These recordings, considered within normal limits by our electrocardiographic criteria, were derived from a population of ambulatory, free-living, and thus presumably normal males. From these data we established the criteria (used in fig. 3) for what is significant change (table 2).

TEMPLATE INVESTIGATIONS

Continuously monitored ECG's taken during real surgery have been transmitted twice weekly for several months on a trial basis from the operating suite of George Washington University Hospital to the MSDL for computer analysis and statistical evaluation. The monitoring system at the hospital serves a cardiovascular operating room, seven general-purpose operating suites, a postanesthesia recovery room, and a special-care unit. These locations contain only the remote monitors such as oscilloscopes and alarm systems, while the signal-conditioning and recording equipment is in a central monitoring room; all signal-routing, recording, and display can be controlled by a nurse at this location.

The data have been transmitted from the operating room to the computer by applying the analog-ECG-amplifier output directly to a portable acoustically coupled dataphone (Bell System-X603C) which fits over the transmitting end of a regular telephone handset. The 603B dataphone receiver at MSDL is then dialed directly over the switched network system. The dataphone transmitter is basically a voltage-controlled, astable multivibrator operating at a center frequency of 1988 Hz. The center frequency is frequency-modulated by the ECG data with a maximum deviation of ±262 Hz. This is well within the passband of the telephone line and can be heard as a high-pitched variable tone at the receiving terminal. Demodulation circuitry in the 603B receiver enables the original ECG to be recovered for presentation to the computer preprocessing system.

Every 90 sec the computer prints six 3.7-sec time blocks of analyzed data processed during that period. It soon becomes quite evident that a large mass of data is piling up; while this accumulation may be desirable from a specific researcher's point of view, the operating-room team would be better served by data-reduction

TABLE 2.—*Criteria for Significant Change*

Item	P	Q	R	S	T	PR	QRS	QT
Duration, sec	0.04	0.02	0.02	0.02	0.06	0.06	0.04	0.08
Amplitude, mV	0.08	0.06	0.60	0.60	0.30			

techniques that would focus on significant changes or simply indicate no change.

The preliminary results of the anesthesia studies showed certain factors to be of interest in reference to template analyses. For example, amplitudes of wave forms have varied over hourly periods by about 50 percent of the true values; duration values have varied by from 25 to 33 percent. It is of interest that sampling of a subject as rarely as every 30 or 40 sec during total anesthesia yields sufficient data for reasonable monitoring. Monitoring during exercise and stress is a separate problem now under investigation.

The results of these studies confirm that any template method depends on the criteria established for significant change. The criteria can be based on analysis of any preselected subject population or time period; it must be tailor-made for the expected use. Very sensitive numerical criteria may indicate "change" so frequently that the user is overwhelmed. Flexible criteria, easily modifiable on request, are more likely to satisfy his needs.

Table 3 shows initial raw data from six subjects monitored on-line in real time by computer from a surgical suite; under anesthesia they underwent uncomplicated surgery, and the data reflect 1 hour of each subject's surgery. Statistical tabulation of the on-line measurements was done off line.

The percentages are the degrees of variation from the means during the 1-hour monitoring. That is, if the mean duration of a wave was 0.04 sec, and if it varied by 0.01 sec, the variation is expressed as 25 percent. The variation reflects biologic and physiologic changes and some stress as well as measurement error; all these factors must be considered in review of any data. When the variation exceeds these limits, the hypothesis is that close scrutiny of the subject is necessary. These values are not out of line with the variations noted between the 2000 subjects analyzed for the template-of-change analysis, previously described. One advantage of computer measurements is that they allow us to define inter- and intra-subject variability with a high degree of accuracy.

Specialized requirements—In certain instances a computer program (or subroutine) must be specially designed to cope with a situation in which data-distortion may be expected because of imposed stresses, or because the subject is engaged in regular activities. Precise study of the PR- and ST-segment regions, for example, is necessary for optimal early detection of abnormal cardiovascular response to stress, or in coronary care. Figure 3 is a facsimile of a printout of a computer program that provides the necessary detail in such circumstances.

The printout of the computer-processed continuous ECG (fig. 3) is for two 6-sec periods, each of which is independently analyzed. Heart rate, elapsed time from the beginning of the recording, and the particular complex analyzed are identified; Q, R, and S amplitudes (A) and duration (D) are calculated. Then follow the detailed measurements of P-R and ST-T segments. The calibration signal and a portion of the analog record corresponding to each time interval also are shown; the complex selected for analysis is identified.

TABLE 3.—*Approximate Mean Values and Variations in Parameter Measurement for Six Subjects*

Parameter	Mean	Variation, %	Parameter	Mean	Variation, %
PA	0.14	66	STO	−0.04	150
PD	0.09	33	STM	−0.02	200+
QA	−0.06	66	TA	0.13	66
RA	0.58	33	TD	0.16	25
RD	0.05	20	QT	0.37	15
SA	−0.19	50	QRS	0.09	20
SD	0.04	25	RATE	75	23

There is considerable interest today in electrocardiographic monitoring for diagnostic testing during exercise, for physical-fitness studies at work-evaluation centers, and for management of acutely ill patients in coronary-care units. These activities place a great burden on electrocardiographers; in response to their needs, electrocardiographic telemetry equipment and monitoring systems have developed rapidly.

Most current monitoring systems employ analog computers which are generally limited to tracking of changes in rate and rhythm of the heart. But during activity the electrocardiographic record may include shifting of the base line, noise artifacts, superposition of wave forms, change in heart rate over a wide range, abnormal conduction, and arrhythmias. Yet the finding that gives the exercise ECG its greatest clinical value, the ischemia-induced wave-form alterations, must be detected and distinguished among the other events. This cannot be accomplished with existing analog equipment.

In our test program, the QRS-complexes in the 4 sec of data are identified by use of minimum derivatives, and the R-R-interval between consecutive beats is measured for detection of cardiac irregularity. A cardiac irregularity is present if in two consecutive R-R-intervals the smaller interval is less than 75 percent of the larger. If there are not two consecutive regular R-R-intervals, the computer printout displays the word "arrhythmia" for that time block, and analysis of the next time period begins.

If an arrhythmia is absent, the program searches by means of a slope check for a region free of base-line shift or noise. This slope-check procedure is applied to paired heart beats in every 6-sec time block until a pair is located that fulfills the criteria for base-line constancy and artifact-exclusion (fig. 4). The program identifies the QRS-onsets of two adjacent cardiac cycles; the line joining these two points is defined as the indicator of base-line slope for these two heart beats. If lines joining other corresponding points elsewhere on the cardiac cycle are found parallel to the indicator line, the base-line slope must be constant for the total duration of these two heart beats. Since time intervals are constant, slopes may be compared simply by comparison of amplitude differences. The computer program uses the amplitude difference between QRS-onsets as the reference for base-line slope. Since physiologic beat-to-beat variation often prevents perfect slope constancy, some discrepancy with the reference amplitude is allowed.

The slope-check procedure eliminates data

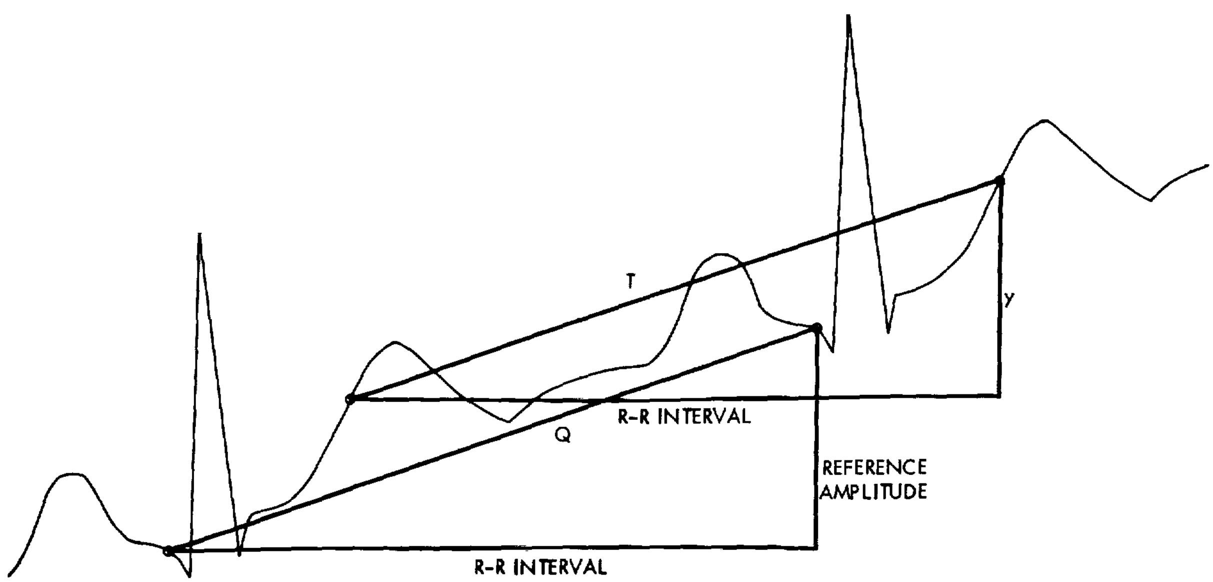


FIGURE 4.—Application of slope-check procedure.

containing excessive base-line fluctuation or transient artifact, and also rejects a region from further analysis if there is disparity with the reference amplitude beyond the established limits. If the slope check fails to find a satisfactory region, the statement "artifact obscures" is printed for that time block.

Once a pair of "matched" heart beats are located, the program proceeds with a measurements routine. All values along the line joining QRS-onsets are computed from the R-R-interval and the amplitude difference between QRS-onsets. Amplitudes are measured as vertical distances from the points on the curve to the indicator line. This technique corrects amplitudes from a sloping base line to a horizontal line; in this way amplitudes along the P-R and ST-T segments are measured relatively to the QRS-onset as the reference level. Commencing with S-T-onset, amplitudes are measured every 20 msec for a time equal to half the R-R-interval. Four amplitudes at 20-msec intervals preceding QRS-onset are measured on the P-R-segment. These measurements and Q, R, and S measurements are printed along with the time elapsed from the start of the recording (fig. 3).

NOISE

The great noise during flight is perhaps the only differentiating feature between clinic and space in requirements for ECG monitoring. Typical clinical noise is shown in figure 5. The amplified version shows actual content; conventionally in medical wards the pen recorder eliminates even large amounts.

Special problems exist in extraction of meaningful data from ECG's monitored during stress, exercise, or certain other conditions. In electrocardiography during exercise we have seen a trend toward continuous recording during the exercise period and increase in the severity of exercise. Although noise can be reduced to some extent by proper application of electrodes and grounding, extraction of representative data from the monitored ECG remains an ever-present problem both for the reading physician and for a computer-analysis program.

Several points need emphasis before discussion of techniques for extraction of meaningful data or reduction of noise in electrocardiographic processing. The ECG is only an index to the electrical activity of the heart. Our understanding of the physiology involved in production of the signal is limited. It is well known that our recording and display devices are often poorly standardized determinants in the representation of the activity observed. Therefore distinction between what is signal and what is noise is necessarily poor.

The physician attempts to make the distinction by relating a particular segment of the signal in question to adjacent segments of the curve in that cycle, and also to corresponding segments in adjacent cycles. He also uses the relations compiled from previous viewing of other curves as well as his knowledge of the physiology responsible for the signal. Distinctions are essentially based on a subjective consideration of probabilities and may not be accurate; for this reason we have suggested increased emphasis on statistical considerations.

A point in the cycle is judged to be noise if it represents an abrupt amplitude difference from adjacent points in the cycle, if it is not repeated (within limits) in a similar position in adjacent cycles, and if it creates a pattern for the cycle that is incompatible with the physician's preconception of a pattern.

For purposes of discussion, noise in electrophysiologic signals may be categorized into base-line shift (low-frequency, cyclic, or noncyclic), cyclic high-frequency (e.g., 60 Hz), relatively isolated spikes, and random noise varying in frequency. Each of these types may exist with various average peak-to-peak amplitude values. In addition, in a condition commonly observed in transitional leads, there are marked changes in waves or segments from beat to beat with no predominant pattern. It is not clear in a given recording whether this condition is due to true variation in the signal or to recording at a particular site.

Several techniques may be applicable to these problems, either singly or in combination, that have been or could be used in both the exercise and NASA's ECG program. As already mentioned we have used two computer routines identified as the amplitude-agreement check (parallel check) and the time-interval-agreement

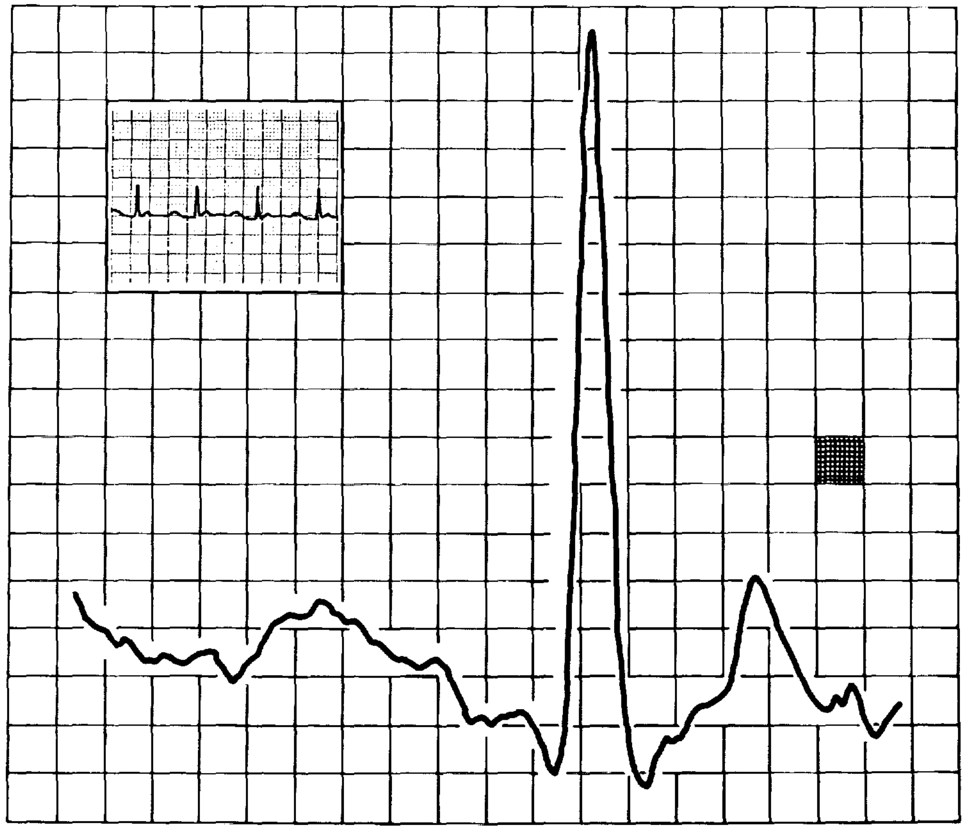


FIGURE 5.—Typical clinical noise.

check; the others are (1) analog filtering, (2) digital-computer smoothing, (3) analog-signal averaging, and (4) digital-computer averaging. These different programs should be evaluated by their utility in extraction of analyzable data from noisy tracings with minimal distortion of signals.

Analog filtering—Analog filtering is the simplest technique for reduction of noise, but it may distort good-quality data by obscuring or falsely producing significant wave forms when applied indiscriminately to the signal. The effect of filtering is, among other things, a function of wave-form, heart-rate, and filter characteristics. There is distortion of the ST-T segment with low-frequency cutoffs higher than 0.05 Hz (at 6 db per octave). The proper high-frequency cutoff for exercise ECG's is less apparent. To show notching and slurring, high-frequency recording fidelity is more likely to be useful in the resting than in the exercise ECG. At present no practical importance can be ascribed to such data. Lowering the high-frequency noise while avoiding, obscuring, or falsely producing diagnostically significant transients raises an important research question. There is still no clear definition of a significant transient. The effectiveness of lowering the high-frequency limit of a filter for production of readable ECG's is also unclear.

Our experience with analog filtering of the exercise ECG, with a bandpass filter allowing frequencies between 0.02 and 45 Hz, indicates that noise remains that complicates our pattern-recognition by computer. Many of these records are judged readable by manual methods. It is of interest to know (1) whether lowering of the limit further (to about 25 Hz) would be effective in producing readable data, and (2) what distortion of good-quality data would accrue at these settings.

Digital-computer smoothing—Several available mathematical techniques may be valuable either alone or in conjunction with other noise-reduction techniques for data. If these techniques are applied indiscriminately to the signal, there is a risk that diagnostically significant characteristics of the signal may be either obscured or falsely produced.

Morphology of the wave forms, heart rate, the type of smoothing technique used, and the number of points involved in the smoothing of a single point are factors that affect both these desirable and undesirable tendencies. Three of these techniques—the moving-point average, the medium-point method, and the parabolic fit by the least-squares method—have been used by us experimentally with the exercise-ECG program. Programs have been written in such a way that the number of points used for each technique may be varied.

Initial experience indicates that the medium-point method works best on data that are distorted by large spikes, of relatively short duration, separated by periods of relatively little noise. Its drawback is the great tendency to eliminate important peaks and nadirs in the data.

The moving-point average is best suited to damping of cyclic noise of high frequency as well as random noise of low average amplitude. It reduces the amplitude of such important peaks and nadirs, but not so readily as does the medium-point technique.

By the least-squares method, the parabolic fit can decrease the amplitude of low-amplitude, high-frequency noise while preserving large-amplitude peaks and nadirs, but it has a tendency to add to peaks if their amplitudes are great. It is less effective than a moving-point average for elimination of high-frequency noise in low-frequency areas, such as ST-T regions, when the same number of points are used for each technique; another drawback is the time required for processing.

Some of the disadvantages of these programmed techniques can be overcome if they are applied only when the data are automatically judged to be of poor quality, and only in specific areas in the cycle that are best suited to use of the particular technique. Further information is needed concerning the effects of these techniques on various regions of the cycle, such as PPR, QRS, and STT regions, when rates and morphologies in these three regions are varied. The techniques should be evaluated in terms of producing some percentage change in the parameters measured, changing of the signal on interpretation from one diagnostic category to another, and the effect on a record's percentage of data deemed processible by the computer program.

Analog signal-averaging—Averaging with a small analog or digital computer can be used. Its limitations are inherent in its continuous effect, inability to deal with nonrandom noise, averaging of homogeneous inhomogeneous transients, and obscuring of diagnostically significant transients. Some of these problems are related to the choice of a fiducial point in the cycle for use in triggering, which is necessary for superimposition of complexes as is done with this technique. The amplitude threshold may shift in time with base-line shift, thereby effecting a requirement for a large number of complexes to be averaged in order to "mean out" the contribution by the base-line shift to the average. Slope triggering has been suggested, but we find no reference to experience with this in analog averaging. Change in rate during the averaging period can also be a problem.

Digital-computer averaging—Averaging with a digital computer has an advantage in that it can be applied to the signal only when the program determines that the data are sufficiently poor in quality to require it. It is possible that the risk of obscuring diagnostically significant transients may be partly avoided by reduction in the number of complexes required for averaging; we have tried this.

The Q-onsets are identified as well as possible in noisy data, and each complex is corrected to the horizontal by use of these Q-onsets even if they are located somewhat in error. Probably it will be necessary to reduce the change and magnitude of the error in location of the Q-wave onset by refining the region of search for its location. Perhaps this may be done by examining a small region about the minimum derivative and identifying the maximum positive derivative in this band. A longer region of search would be necessary for instances in which the maximum positive derivative was located to the left of the minimum derivative and vice versa. After base-line correction is accomplished, complexes can be superimposed by alignment of minimum derivatives.

The sampling of a complex to be averaged would end at a point located after the minimum derivative, at a distance of 75 percent of the interval between minimum derivatives. A point located before the minimum derivative, at 50 percent of the distance between adjacent minimum derivatives, would limit the sampling region at the beginning of the complex. It will be noted that with this method there would be some overlapping of data, with the same data appearing on both ends of the complex being averaged. Data will be eliminated from the average, from both ends of the averaged complex, beyond the point at which the shortest cycle corresponds in time to the other cycles. The effect of base-line shift may be minimized by this method, but the method shares other disadvantages previously mentioned of the analog averaging technique.

This technique, however, has an advantage over filtering and smoothing techniques when there is large beat-to-beat variation or when noise-deflection durations approximate the length of recognizable segments and waves of interest. Both the smoothing and filtering techniques reduce noise by relating displaced points to other points in the region, and cannot be expected to be effective with problems such as beat-to-beat variation in which all points in a region or segment may be displaced. Averaging has an advantage in this situation since with it a "displace" point is related to a corresponding point in other adjacent complexes.

Noise and satellite-data transmission—To demonstrate the feasibility of sending signals over very long distances for on-line computer analysis, with subsequent return of the interpretation, a trial communications system was set up between Tours, France, and MSDL in Washington, D.C., making use of the Early Bird satellite. Both 12-lead resting and continuously monitored ECG's were transmitted for almost 2 hours daily for 5 days (3–8 July 1967) with signals comparable in quality to local transmissions.

The ECG signal was conditioned with a portable ECG-encoder for patient identification and acoustically coupled to the French telephone network for transmission to the Comsat transmitter at Pleumeur Bodou, France. The acoustically coupled FM signal was then used to frequency-modulate a Comsat carrier frequency for transmission to Early Bird which resides in a synchronous orbit 22 236 miles from Earth. The signal was retransmitted to the receiving station at Andover, Maine. At that point the original frequency-modulated 2-kHz tone of the acoustic coupler was recovered and placed on the private telephone lines of RCA Communications for routing to Washington. When the signal reached Washington, it was transferred to domestic lines for transmission to the Bell System 603B dataphone receiver at the MSDL.

The signal was filtered, amplified, and converted from analog to digital form at a rate of 500 samples per second for presentation to our digital computer (CDC 160–A) for measurement and analysis of the various amplitudes and durations. The computer produced punched paper tape that was fed into a teletypewriter (Bell System model-35 ASR) for transmission to New York where it was necessary to convert from the domestic eight-level code to five-level International Telex code. The speeds of the machines associated with these codes are 100 and 60 words per minute, respectively. The information loop was closed when the ECG interpretation was returned to Tours via the Early Bird satellite and Telex lines.

The quality of the signals received in Washington was very good. The only noise showing on the paper tracings was that due to occasional muscle movement or electrode slippage. The first attempt at transmission was greatly affected by echoes; at one time 12 echoes were audible after a spoken word. This fault was soon corrected by insertion of echo-suppression devices in the line at the RCA terminal. During transmission it was also necessary to disconnect the telephone receiver in France because the original transmitted signal was returning, after a round trip of approximately 0.6 sec, completely out of phase with the real-time signal. In addition to the cross talk created on the lines, the signal was being acoustically coupled back into the transmission circuit. Insulation of the acoustic coupler from room sounds eliminated the ambient noise. No difficulty was experienced with the teletypewriter interface, and the transition from 100 to 60 words per minute was smooth.

Such a system is obviously expensive, but in special situations the returns would be commensurate. First is the use of satellite transmission for monitoring of astronauts. Underdeveloped nations, lacking funds or technology, might benefit from the far-reaching implications of such a system to their national health. Even more obvious is the potential of medical-signal analysis or storage centers within countries like the United States with existing communications facilities.

PREDICTIVE VALUES

Rapid statistical analysis is possible during monitoring if data are in digital form. From such data, statistical displays can be obtained on line and in real time to provide, for example, a standard score value for each parameter (or a multidimensional score for all parameters) or a status report on a subject for comparison with his past results (or for his comparison with a specified population group). By these and similar measures, a subject's profile can be quantitatively displayed.

Statistical tabulation and analysis of data are intended to provide insight into the general, central, or predictive trends of the parameters being measured. Heart-rate change can show how data can be used for predictive purposes. The heart rate at rest is generally used as the fiducial point in stress, exercise, or continuous monitoring, or simply for pulse-determinations; the generally accepted normal range is from 60 to 100 beats per minute.

Analysis of 27 000 computer-measured ECG's has allowed a better distinction of usual heart rates and those that could be beyond usual limits. Our data make it apparent that two considerations are necessary for judgment of heart rate. First, a subject's rate should be within certain defined limits when at rest. The second consideration is of particular interest; our large accumulation of data has allowed insight into variation between individuals and between disease groups even during rest. Thus an individual's variation in rate must also not exceed the variability within his specific category of physical fitness or disease, or that of the "normal" population's distribution.

Application of numerical values—If the values of several astronauts on several different flights were surveyed, different functional time periods or different types of stress would be studied. It would then be useful to compare quantitatively the responses of new astronauts in training with those of astronauts in flight. One result would be a more quantitative basis for selection of trainees.

During flight the results for the total period from one tracking station could be used to establish a trend during the flight for contrast with preflight information obtained from test chambers. On-board ECG tapes or tracking-station recordings could be evaluated for statistically meaningful data relating measured parameters to space-flight activities. From the research viewpoint we could quantitatively study changes associated with weightlessness, particularly during ascent, orbiting, space walks, reentry, and recovery—in fact during any conditions that cannot be tested on Earth's surface.

To demonstrate such application of numerical values, data have been compiled from on-line computer measurements of the pilots in Gemini flights 9 to 12. The easiest parameter to understand is heart rate; it may not be the best from which to draw conclusions about status of the subject, but it is a reasonable one to consider first. Figure 6 shows the heart rates of subjects during different flights. Compare, for example, Gemini 11 with flight 12 or 10. On flight 11, during the upper quartile of the time, both pilots had heart rates of about 90 to 100; this did not occur in the other flights. Their average heart rates also were higher than those of pilots during other flights.

The heart rate was therefore variable, depending on the number of the flight (i.e., problems) as well as on each individual. At the 75-percentile and the 50-percentile levels, for example, either flight 10 was less stressful or perhaps the pilots were better trained than for flight 11. Whether the variation is due to functional activity, individual physiologic response, or undetermined causes, it is important that levels be determined in training and in flight, for each individual pilot or mission, to form the basis of a template for comparisons of one pilot with others. Subsequently the levels could serve to establish standards for pilots in varying types of flight.

As we have mentioned, heart rate is only one of the parameters of a subject. The availability of other computer-derived data makes of paramount importance definition of the quantitative

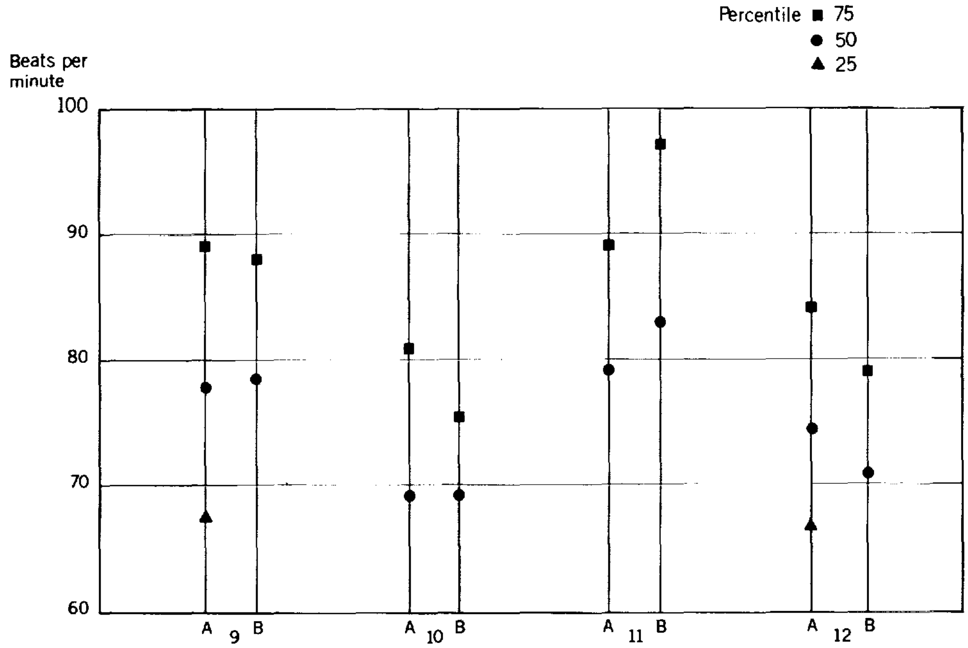


FIGURE 6.—Selected percentile of heart rate.

usefulness of those data in determination of the status of the subject.

Obviously many limitations are present in these crude graphs since for this analysis we are not considering functions being performed, time of day, etc. We are just taking data computed for the total flight, but these are useful in showing that a trainee's reactions could be contrasted with this total experience. It is also useful to consider that each pilot in each subsequent flight could be contrasted in real time with the total experience of all previous flights. The difference between one pilot's experience and the total experience or that of a similar series of flights—the trend to or away from the total experience—might be significant (fig. 7).

The data in reference to other measures may be of equal importance. Consider the QT-duration shown in figure 8. Although there is similarity between various subjects, we could say that subject B on flight 11 had a lower QT than others. We could compare that fact with heart rate (fig. 6) and note that the heart rate was higher. This subject could be assumed to be different because of his tasks, physical attributes, or environmental circumstances. The cumulative percentage of QT's, had they been available for all previous flights as for 9 to 12 in figure 9, would have shown by how much the subject deviated in performance from all previous flights, preflight experiences, early flight stages, or from all other astronauts. Contrasts can also be made with the general population or other groups. This illustrates the fact that instantaneous, multidimensional, on-line, statistical analysis can give useful indications of each subject's performance.

Another parameter studied was QRS-amplitude which is complicated by the fact that NASA's electrode placements tend to cause rapid changes in polarity; but, even with this artifact, differences existed between flights 9 and 12 and 10 and 11 (fig. 10). No conclusions can be drawn from these data; justification for further study is apparent. The onset of the ST-segment shows similarity among subjects 10a (high quartile) and 12a (low quartile) that may be due largely to computer-measurement and telemetric problems, but this otherwise suggests that a range of ST-onset can be defined. Under known conditions a subject should not vary too much from this range (fig. 11). The STM (midpoint of the ST)

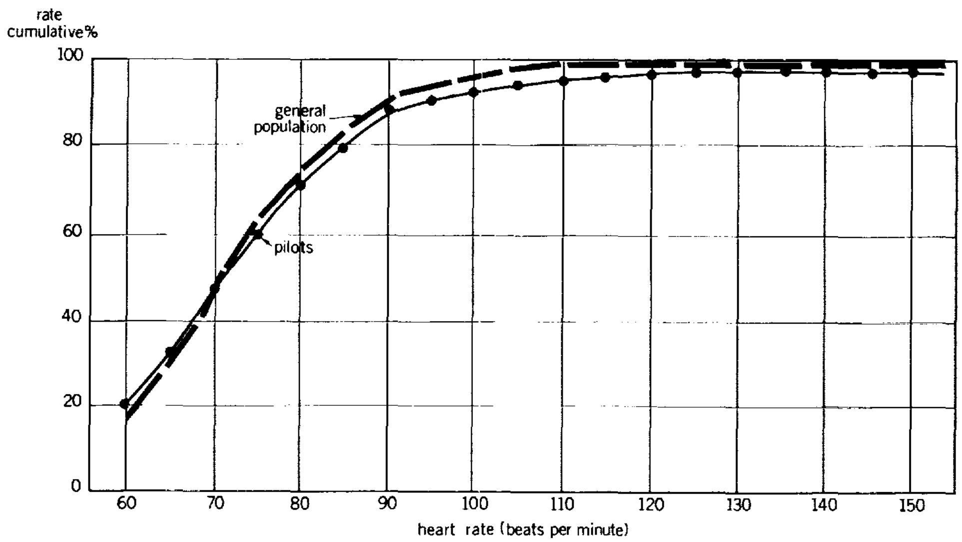


FIGURE 7.—Cumulative percentage distribution of heart rate.

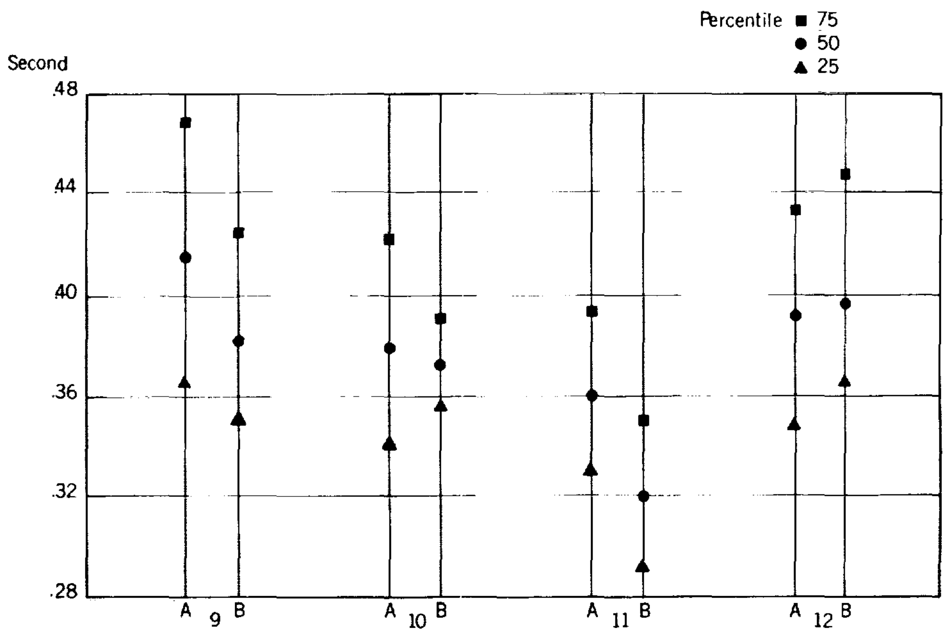


FIGURE 8.—Selected percentiles of QT.

was beset by noise in our signals, but with more data its range could become significant.

The PR-interval is an area that shows wide variation (fig. 12). This was the area of greatest noise in the NASA tracings, as received by us, so that no conclusion can be made. Since this area is an early predictor in clinical medicine, it could be useful in space medicine and warrants special techniques for noise reduction on line.

It is apparent that functions of the individual and noise in the signal interfere with the statistical analysis to be used. The noise has been indicated to be of various types, and suggestions for certain of the areas of improvement have

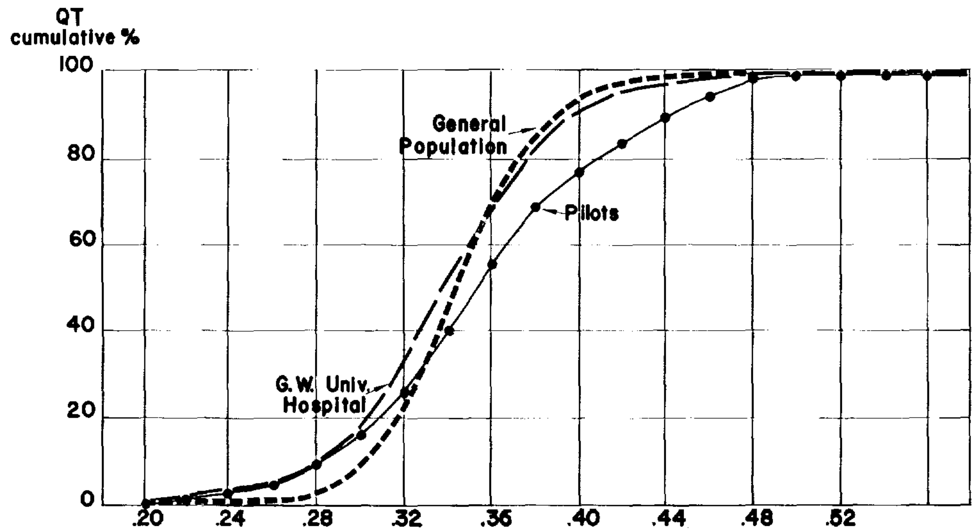


FIGURE 9.—Cumulative percentage distribution of QT.

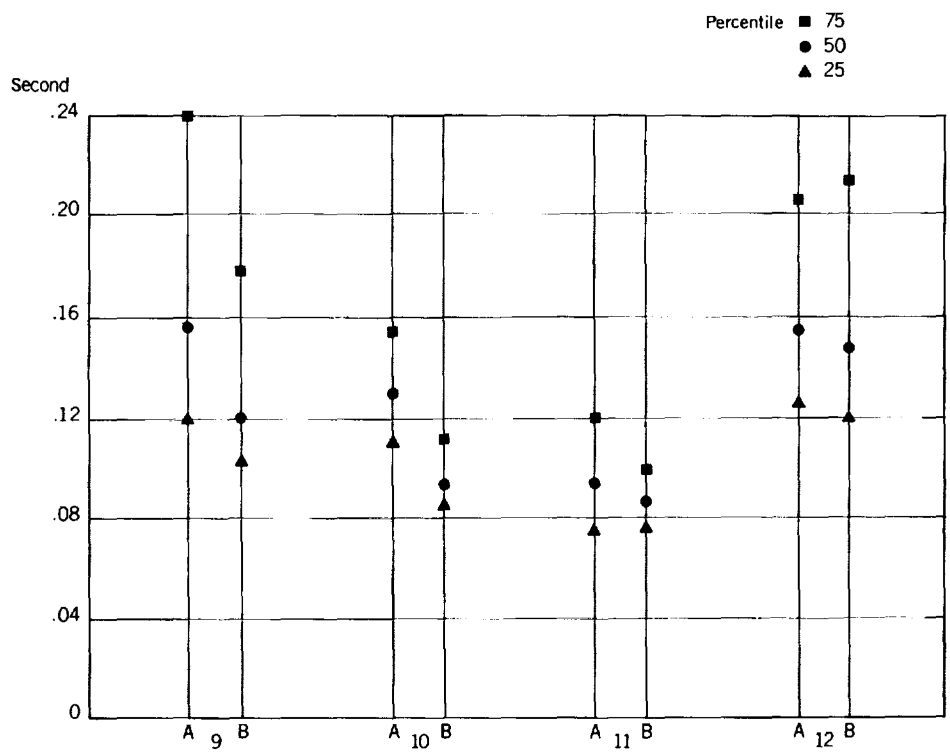


FIGURE 10.—Selected percentiles of QRS.

been noted. In our experimental transmission from Tours, the signals during activity were remarkably free of noise. It would appear then that much of the problem of noise can be solved by satellite rather than by systems currently in use. The suggestion is that this change should be made as rapidly as possible so that the statistical techniques can be used and full advantage taken

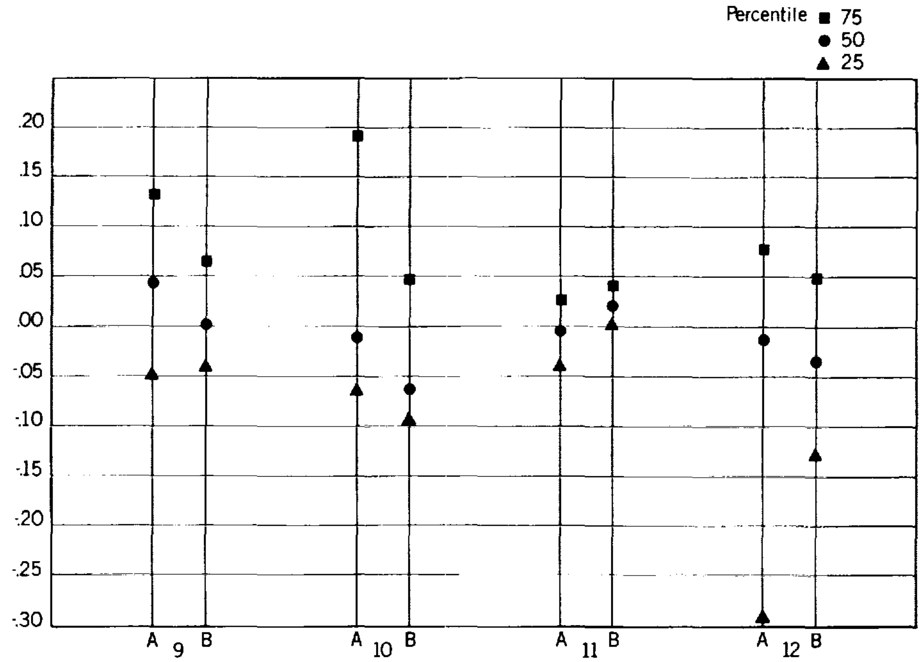


FIGURE 11.—Selected percentiles of ST-onset.

of the possibility of measurements by computer.

There are several preliminary recommendations on the basis of the Gemini flight data. First, there is need to study preflight training information and to compare it with general-population and other data to determine wherein special templates, if any, are required for astronauts. This recommendation extends to all students of special groups such as those in intensive-care suites. Second, it would be reasonable to determine better ways of pilot-selection by means of quantitative data. Third, and perhaps most important, the functional capability of pilots could be complemented by knowledge of each other's reactions, in terms of quantitative variables, under various conditions of stress and activity.

Statistical potential—Continuous monitoring of the ECG is needed to provide useful clinical, physiologic, and epidemiologic data, during exercise or stress tests, in the management of patients in intensive-care units, in operating and recovery rooms, and in the routine follow-up evaluation of heart patients. With the present empiric system of monitoring, the ECG cannot be fully used for prompt detection of early changes of significant magnitude. Three factors are responsible for this:

(1) Empirically there is no way to analyze all the values of amplitude and duration of the continuous electrocardiographic wave forms. Because of the nature of biologic variability, spot checks or sampling methods of interpretation are also generally unacceptable. Thus only crude measures of heart rate, heart rhythm, and an occasional wave form are possible by conventional methods.

(2) When the data are finally available to the monitoring personnel, it is usually long after the fact. Thus a situation exists in which the information reaching the physician is either insufficient or too late.

(3) Little effort has gone into study of evaluative procedures.

One possible procedure to overcome these limitations is establishment of statistical computer programs that can be used to "format" displays of immediate verbal statements classifying the measurement patterns into clinically diagnostic or significant categories. These programs can take the electrocardiographic measurements for selected time segments and relate them to previous data.

The purpose of this section is to present as an example a statistical method for determining

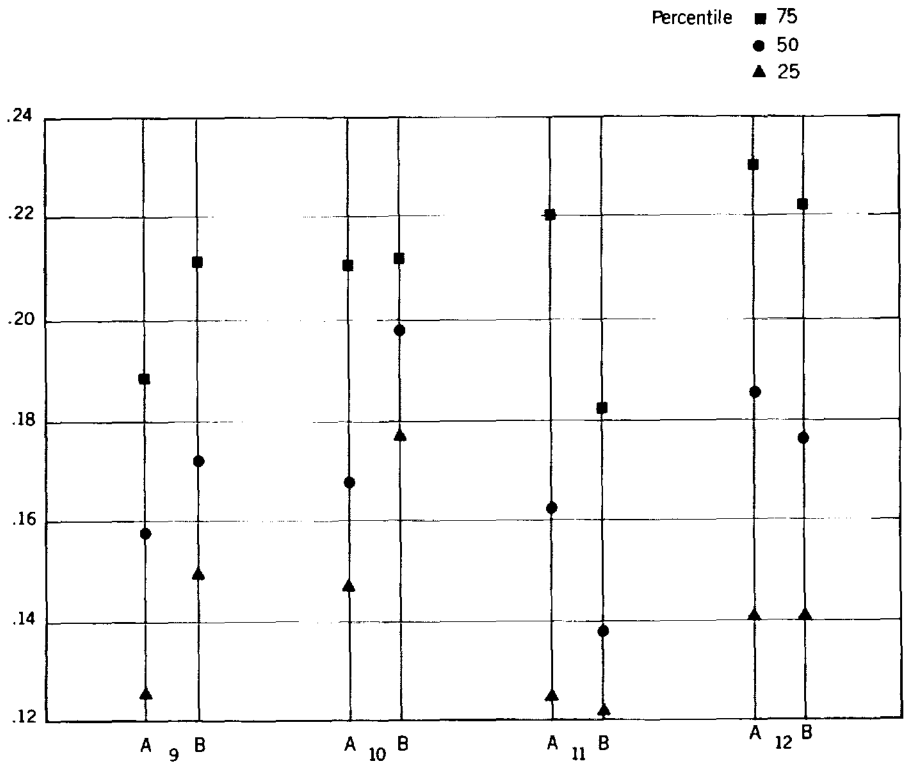


Figure 12.—Selected percentiles of PR-intervals.

significant change in the continuous monitoring of the ECG. The method selected is Hotelling's T^2-multivariate technique; it requires only the extraction of information from any large pool of data such as that derived from our ECG computer program.

The conventional ECG is a 12-lead signal with approximately 25 measurements of amplitude and duration per lead; the signal is processed by computer to yield amplitude (A) and duration (D) of individual waves. The time axis is the independent parameter; amplitude or height, the ordinate of the curves, is the dependent variable.

The 12-lead ECG can be represented as a column vector of 300 variables. Information is potentially present in each of the individual variables and also in combinations of variables from crossing leads, such as lead-1 and lead-3. Many of the diagnoses are based on information from both sources—individual variables and combination variables—but all possible crossover, intra-inter, lead configurations of parametric measurements total at least 300^2 compared to only about 100 possible diagnostic-statement categories. This implies that in fact only a small fraction of all possible information is being used for current diagnoses. One reason that can be advanced for such nonutilization of all the data is that no convenient method with easy calculation exists for extraction of the additional information. It is in this area of electrocardiography that its future utility may lie.

Statistical-model project—We have selected electrocardiographic measures for input to a T^2 program that uses these data to build a pool of predata estimates. From the predata one computes a premean column vector and a prevariance-covariance matrix. If one assumes that the electrocardiographic signals are jointly distributed as multivariate normal with a mean (μ) and covariance sigma, then the quantity

$$N(\bar{x}-\mu)'S^{-1}(\bar{x}-\mu)$$

is distributed as T^2 where N is the number of samples used to accumulate the prepool from which $\bar{x}$ and S were calculated. For any post-sample pool of size N_2, the quantity

$$\frac{N_1 N_2}{N_1+N_2}(\bar{x}_1-\bar{x}_2)'S^{-1}(\bar{x}_1-\bar{x}_2)$$

is distributed as T^2 with N_1+N_2-2 degrees of freedom where S is now the pooled covariance matrix for both states, pre and post.

For any signal taken during an operation, a flight in space, or in a recovery room, a statistical comparison can be made with the prepool data by substituting for $\bar{x}_2$ in the expression with $N_2=1$. That is, for any electrocardiographic signal, x_0, one computes

$$\frac{N_1}{N_1+1}(\bar{x}_1-x_0)'S^{-1}(\bar{x}_1-x_0)$$

The T^2 value can be determined for each designated block of time in the post period. For any level alpha (α), the critical region for T^2 is

$$T^2 \geq \frac{(N_1+N_2-2)p}{N_1+N_2-p-1} F_{p,N_1+N_2-p-1}(\alpha)$$

where p is the number of parameters considered in the electrocardiographic analysis. If the T^2 value exceeds the critical value for the 5-percent level of significance, the statement is made that the current sample is significantly different from the prepool estimate. To the physician, nurse, or ground crew monitoring the subject's ECG, a significant T^2 value indicates that the signal is statistically different from the electrocardiographic prior data. The relation of statistical significance must of course be determined by experimental studies.

Resting electrocardiogram—Twenty electrocardiographic measurements from 59 subjects were used for the predata estimate. The ECG's were diagnosed as within normal limits both by the computer program and by two physicians who used the standard 12-lead ECG as the source of information. Twenty electrocardiographic measurements from each of 10 subjects, seven within normal limits and three beyond, were then compared with the prior-population figures. The measurements of each of the 10 subjects, known to be either within or beyond normal limits on the basis of standard 12-lead ECG's, were obtained from 3.7-sec periods of continuous electrocardiographic signal, so that we have simulated a T^2 evaluation of 10 separate column vectors of electrocardiographic data on a continuous-monitoring basis (table 4).

TABLE 4.—*Classification of 10 Patients by the T^2-Technique*

Classification	Normal limits	
	Within	Beyond
Correct	6	2
Incorrect	1	1

Of the three ECG's that were not within normal limits, two had significant T^2 values or a correct classification. Of the seven ECG's within normal limits, six were correctly classified. Overall the T^2 test showed 80-percent agreement with the standard 12-lead ECG.

The T^2 test used 20 electrocardiographic measurements for evaluation of the continuous ECG. The measurements were few relative to the number used in the standard 12-lead ECG for evaluation. It is of interest to compare these findings with the findings of five independent readers of the electrocardiographic tracings from which the 20 measurements were taken (table 5).

Of the six ECG's within normal limits by both the standard 12-lead ECG and the T^2 test, all readers agreed with this evaluation on four. At least two readers judged the remaining ECG's beyond normal limits. Of the ECG's beyond normal limits by both methods of evaluation, all readers agreed with both methods on one; at least one reader similarly agreed regarding the other.

Regarding two ECG's there was disagreement between the standard 12-lead ECG and the T^2 test. On one ECG, at least one reader agreed with the standard method. On the remaining ECG, at least one reader agreed with the T^2 test. In summary, at least one of the readers, using the same limb leads from which the electrocardiographic measurements were taken for the T^2 test, agreed with the statistical method in each case.

Exercise ECG—Electrocardiograms were obtained from one subject at rest, during exercise,

TABLE 5.—*Electrocardiographic Findings by Five Readers for Subjects on Which Standard 12-Lead ECG'S and T^2 Test Agree and Disagree in Evaluation*

Finding	ECG and T^2			
	Agreement		Disagreement	
	Within normal limits	Beyond normal limits	ECG normal; T^2 abnormal	T^2 normal; ECG abnormal
	All readers agreed			
Within normal limits	4	0	0	0
Beyond normal limits	0	1	0	0
	At least one reader agreed with T^2			
	2	1	1	1

and after exercise (the recovery period). Twenty-two column vectors of electrocardiographic data were recorded from the subject seated at rest on a bicycle; after exercising long enough for his heart rate to reach 150 beats per minute, he rested. Electrocardiographic recordings continued until the subject's heart rate returned to within 10 beats per minute of his resting rate. There were four blocks of exercise recording and 27 blocks of recovery-time recording.

For the T^2 test, the 22 time blocks of resting ECG's were used as the predata estimate; each block consisted of 20 electrocardiographic measures of interest to the physician. The variance-covariance matrix and the vector means were calculated from the 22 time blocks of data. The exercise and recovery column vectors were compared with the predata estimate.

The T^2 value was significant for the first, third, and fourth exercise time blocks, but not for the second. The first block of the recovery period was significant; the remaining periods were not. The program could demonstrate at what time a subject returned to resting level and differentiate the exercise period from the recovery period.

Perspective—There are many possible ways to monitor the measurements of variables from a continuous electrocardiographic wave form. One approach is by detailed comparison of these with measurements that have been considered standard. Another approach is to formulate values to relate measurements in a defined way similar to that in which an electrocardiographer or a text book would relate them. This method considers magnitude and direction of the wave form; the interpretation of specific patterns is essentially empirical, resulting from clinical and autopsy association. These two approaches are commonly used in clinical practice. A third approach, that of classic statistics (as described), is by statistical study of the joint probability distribution of the electrocardiographic observations. The distribution in a prior population is described by means, variances, and covariances and related to observations in succeeding time blocks of electrocardiographic data.

It is apparent from the review of the limited quantity of electrocardiographic data (20 variables) that not all data usually available to the electrocardiographer, from his inspection of 12-lead standard ECG's, were used in the T^2 test. We were limited to 20 variables because of the computer's limited storage. The absence of an on-line computer program for automatic retrieval of electrocardiographic data also limited the number of samples for building of a prior matrix of measurements.

We have not yet demonstrated to our satisfaction that smaller quantities of data than are generally used are all that are needed to minimize false positives and false negatives in successive time blocks of observed data. The monitoring system, however, fills the need and offers many advantages over the present means for analysis of the monitored ECG. The method reduces a mass of data and provides the physician with clinical information about important wave-form changes; results are stored and retrieved easily for rapid statistical analysis. The method achieves its purpose by supplementing clinical judgment. The additional information gained from specific time periods aids final evaluation of the ECG.

A statistical technique combined with a high-speed computer provides a promising method for evaluation of continuous monitoring of the ECG. This combination offers the physician supplementary clinical information to improve the diagnostic process.

AUTOMATIC CARE AND EVALUATION UNITS

There are many clinical spin-offs from this project. The principles and techniques used in monitoring astronauts in flight can be used effectively for patients in a coronary-care unit or undergoing surgery. Indeed our preliminary tests indicate that both these uses are entirely feasible even though the techniques must be adapted to a hospital environment. The first project to put these monitoring techniques to clinical use is in the testing stage in an operating room and the coronary-care unit (CCU) at George Washington University Hospital.

The nurse and physician in the CCU must now evaluate the outputs of multiple electronic monitors, as well as frequent recordings of pulse, blood pressure, respiration, temperature, urine output, and general status of the patient. With the advent of other automatic constant-monitoring devices, such as for EEG's, respiration, skin temperature, and cardiac output, the medical team will be faced with a bewildering display of data, making their monitoring and decision function more difficult and retarding crucial time-dependent therapy.

The medical team in a CCU thus devotes much time to observation, recording, and analysis of repetitive data. Monitored data on ECG's, blood pressures, pulses, respirations, and other factors are tabulated, correlated, and interpreted. In so doing the medical team acts as a data-reduction system. All these tasks are well suited and easily amenable to digital-computer manipulation.

We suggest development of computer programs and a hardware system for the tasks of a CCU in collection, reduction, analysis, and interpretation of medical data. A computer can be programmed to accept, through standard commercial input-output devices, the routine data that the nurse and physician usually enter in the chart. Data from the clinical laboratory also can be entered. This computer can be linked to another that is able to analyze medical signals such as ECG's and EEG's in real time. This system may be applied to any intensive-care situation such as in recovery rooms and intensive-care units for stroke.

The complex interrelations between clinical data, transducer data, and laboratory data can be evaluated by the computer at each new data entry. All data stored for defined past periods can be reexamined in the light of the new data, and the computer can reduce the data to an English statement such as "stable," "hypokalemia by ECG-check blood chemistry," or "impending shock." This information can be immediately available on a screen in the CCU.

The promptness of therapy after onset of arrhythmias, cardiac arrest, and shock has been shown to be directly associated with therapeutic success. This system allows rapid interpretation of these states even when no physician is present, shortening the time lag between interpretation and initiation of therapy.

Furthermore the system will perform the nurse's routine evaluation, and charting of medical data will supply her more readily with additional information when required; it supports her observations by monitoring certain aspects itself (ECG, EEG, and blood pressure). Thus the nurse will have more time for care of patients and for complex techniques. The immediate entry into computer memory of drug administration and procedures will allow the computer to double-check all new orders for conflict, acting as a safety check to prevent overdosage or use of inappropriate drugs.

Figure 13 shows the output of the initial monitoring program as used in a hospital ward, and a sample of the ECG that was monitored. In this printout each of the columns labeled 1 to 6 is one time block of 4.8 sec. In each block one typical wave form was picked and thoroughly measured. As with our routine program, the whole block was scanned for indication of arrhythmia. Certain problems have arisen in the use of this program, but the direction of the solutions is

INSTRUMENTATION FIELD STATION HEART DISEASE CONTROL PROGRAM
ECG MONITORING

DATE 09-13 ET--19-07-41 ONLINE G. W. OPERATING ROOM MONITORING

TIME	1	2	3	4	5	6
PA	.34	.34	.34	.35	.34	.35
PD	.10	.15	.13	.10	.15	.14
UA	-.23	-.23	-.20	-.22	-.23	-.22
UD	.03	.02	.03	.02	.03	.02
RA	.87	.98	.89	.03	.95	.90
RD	.02	.02	.02	.03	.02	.02
SA	-.42	-.39	-.65	-.39	-.46	-.40
SD	.06	.07	.07	.06	.06	.06
ST	.04	.03	.08	.02	.03	.03
STO	.33	.37	.36	.31	.29	.33
STM	.36	.37	.40	.33	.29	.34
STE	.37	.36	.44	.33	.30	.35
TA	.44	.32	.14	.31	.33	.36
TD	.14	.14	.09	.16	.15	.15
T'A	-.23	-.24	-.21	-.21	-.24	-.22
T'D	.09	.09	.09	.08	.09	.10
PR	.26	.26	.29	.26	.25	.26
QRS	.11	.11	.12	.11	.11	.10
QT	.39	.38	.39	.38	.39	.39
RATE	79	78	77	76	76	76

ST ELEVATION; RULE OUT EARLY REPOLARIZATION
EXTREME ST DISPLACEMENT
TALL P WAVES, P - PULMONALE
PROLONGEC PR INTERVAL, FIRST DEGREE ATRIOVENTRICULAR BLOCK

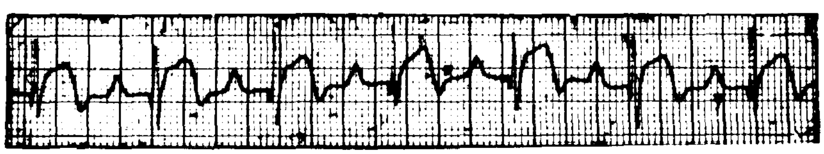


FIGURE 13.—Output of the initial monitoring program and a sample of the ECG.

clear and in progress. One problem is categorizing of complex arrhythmias; another is the variation in calibration in the monitoring signal. The designers of CCU equipment have followed NASA's precedents and have not used clinically conventional calibration routines. Calibration is set by the nurse so that the wave form is large enough to trip the rate meter and therefore be counted accurately. If we are to rely on measurement of the wave forms for statistical purposes, we should introduce a calibrated signal into the system.

Logic flow for surveillance systems—A flow diagram has been developed that outlines the system that we are designing (fig. 14). The system mimics what the physician does, but supplements his logic with statistical techniques in order to arrive at its conclusion.

In the flow diagrams (fig. 14) the ovals enclose statements produced by the system on the screen or typewriter; the rectangles enclose various logical operations that are to be performed and are not detailed in the flow diagram; and the hexagon denotes values calculated from known

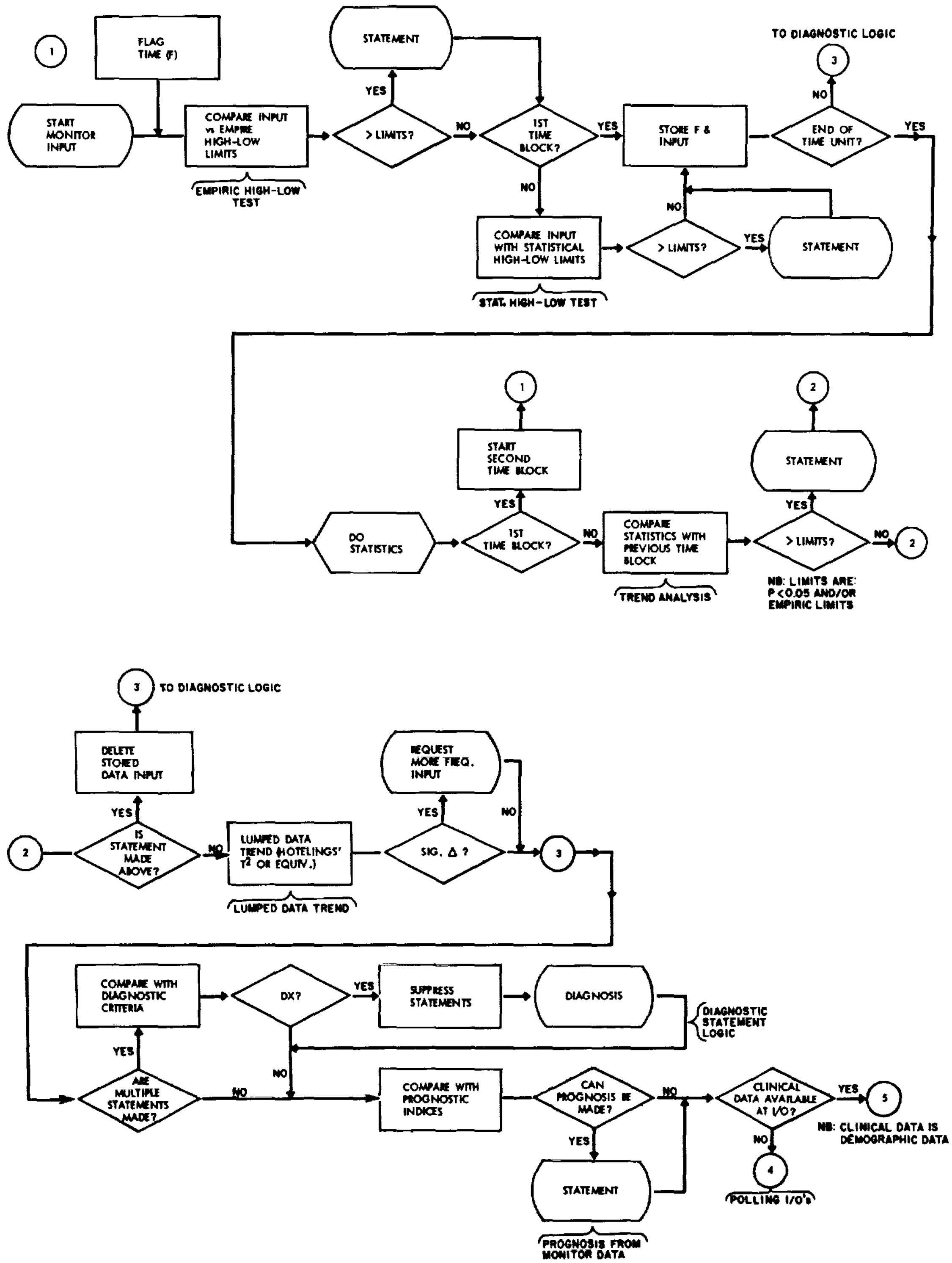


FIGURE 14.—Logic flow for CCU surveillance: flow diagrams.

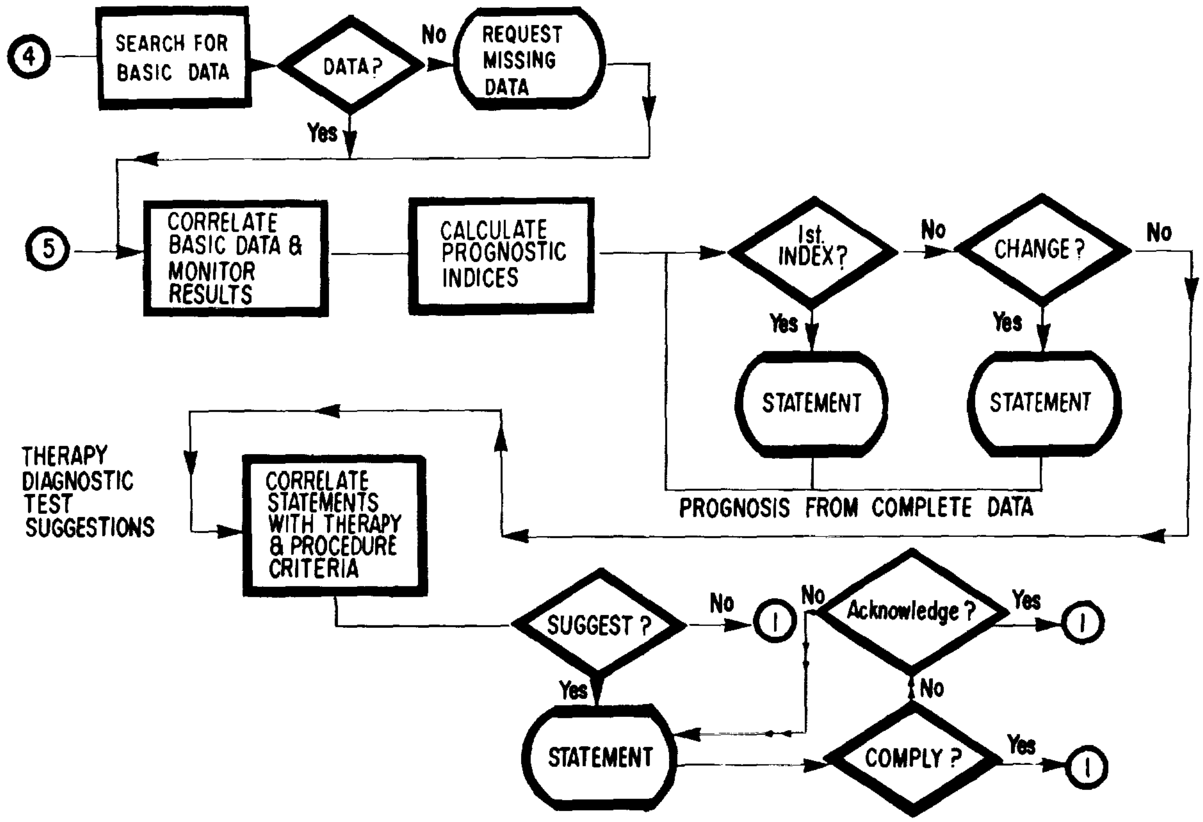


FIGURE 14 (Concluded).—Logic flow for CCU surveillance: flow diagrams.

formulas. The diamond-shaped boxes indicate decisions to be made by the computer, always of the "yes-no" variety. The circles indicate areas that are to be continued on other parts of the flow diagram. The line connecting various boxes in the flow diagrams is to be considered a line of action of event. In order to follow any particular data input one should start at the top of the flow diagram and follow one set of answers; the diagram should lead him sooner or later back to item-1.

The data are first compared to empirical high-low-limit tests to see whether they exceed these empirical limits. Then statistics are calculated from the data for a period of time, and data are compared to statistically derived high-low-limit tests for the past time period. This procedure mimics what the physician does when he says that particular values are unusual for a given patient.

These statistics can then be used for calculation of trends, that is, to see whether any of the variables are changing at a rate higher than is acceptable. Trends are calculated by comparing means and standard deviations of a time period with those of the previous period and noting whether there is significant difference either empirically or on a probability basis. This is equivalent to the physician's noting that the acceleration of the heart rate is greater than normal. If no abnormality is found, the data can be passed to the so-called lumped-data trend-analysis, a statistical method for mimicking the intuitive feeling that something is not quite correct in the mass of data. Such tests as Hotelling's T^2 multivariate analysis can be used to show that something has varied within the data, but they cannot pinpoint the item that has varied; they are more reliable than intuition, however, and are positive when a group of variables begins to change, although none of the individual variables may have changed significantly. Comparison of the monitored data with various prognostic indices can lead to a decision as to whether a prognostic statement can be made.

COMMENTARY

With enough data and experience we will be able to evaluate new relations and to begin, by statistical techniques such as multivariant analysis, to sharpen our ability to predict clinical phenomena such as shock or arrhythmias.

Application of continuous monitoring of ECG's to postinfarction patients, or during anesthesia and surgery, can illustrate its value. The method can also be used in the laboratory, for example, to assess the effects of stress and the effects of radiation on the cardiovascular systems of animals (as in the production of arrhythmias).

In clinical medicine, patients are intensively monitored in the operating room, the obstetrical suite, and other specialized intensive-care units; the procedures are not essentially different from those used in space flight. Personnel in these units perform tests and intuitively obtain results without the objectivity of statistical procedures. In the clinical setting, in addition to natural human bias, the major current disadvantage of functioning with a human monitoring system is simply the lack of trained personnel to perform all the needed services.

The problems of unavailable personnel and reduction in bias of existing personnel can be solved by automation. Furthermore, automated care-evaluation units can serve throughout a hospital, not simply the needs of specialized units. The automated care and evaluation unit can provide for the acute as well as routine phases of health services. Indeed these two phases are not distinct since the patient may quickly change from one status to another. Its ability to perform either phase implies that the unit can carry on the other. The same concepts can hold for space flight. Monitoring multidimensionally and comparison of the results with those from selected population samples can, for example, be the key to proper selection of personnel for space flight or to selection of medication.

For the modern monitoring system to accomplish these goals, three separate items of hardware are required. The first includes transducers and data-acquisition devices; modification, improvement, and development are much needed in this field. The second incorporates computer programs for routine control functions and rapid data analysis. The third includes telemetric data-communication systems to enable easy storage, retrieval, and display.

Immediate association of data is necessary for insight into the current status of subjects and for physiological research to obtain more knowledge for the improvement of care. We envision automated care and evaluation units providing for analysis of many body functions, including ECG's, heart sounds, cerebral electrical activity, vital signs, and vascular, respiratory, and metabolic conditions. Chemical analysis also must be incorporated, integrated, and related to other variables. Thus the automated care and evaluation unit should have facilities for on-line, real-time statistical analysis of data, which should be in a convenient format for use in real time for patient care.

These automated care and evaluation adjuncts must serve and not supplant the human monitor. The automated care unit can and should perform the tasks that are at best routine and noncreative.

SUMMARY AND CONCLUSIONS

Constant monitoring of subjects and data interpretation for evaluation or care are often humanly impossible because the data accumulate faster than they can be analyzed. Use of modern computer systems and statistical techniques allows a new dimension in the quality of medical care that the physician can give.

CHAPTER 8

PERIOD ANALYSIS OF AN ELECTROENCEPHALOGRAM FROM AN ORBITING COMMAND PILOT

Neil R. Burch, Ronald G. Dossett, Abbie L. Vorderman, and Boyd K. Lester

Recording of an electroencephalogram (EEG) from a pilot in orbital flight offers an unprecedented opportunity to inquire into neurophysiological-behavioral relations during a situation unique in recorded history. While the opportunity is unprecedented, this very fact makes full interpretation of the data from any biological system extremely difficult. The single-channel or double-channel EEG, as a psychophysiological measure, requires a great deal of information about the stimulus field for optimum interpretation of neurophysiological-behavioral relations; but by the very nature of the flight situation the stimulus field is impossible to control, and the stimulus field of ongoing events is difficult to chronicle with a high degree of resolution in time.

In the absence of stimulus-field information the EEG is most helpful in quantitative determination of the state of consciousness of pilots in orbital flight. The results of analysis of over 50 hours of continual recording confirm that sleep during this orbital flight was indeed disturbed in a way that could not have been predicted from laboratory experiments, although most individuals show a somewhat-disturbed sleep pattern on the first sleep night in even a moderately unusual environment such as a strange room and bed (ref. 1). At the other end of the state-of-consciousness spectrum (ref. 2), the arousal of alertness and even hyperalertness may be expected to result from a situation as novel as space flight, so novel that it had then been experienced by only several dozen humans.

In interpretation of these EEG tracings several other laboratories have employed a number of different analytical techniques (refs. 3 and 4); even a truncated version of period analysis has been undertaken (ref. 5). Clinical interpretation of portions of this recording has been reported (ref. 6) and may be considered the "standard" system for interpretation of states of consciousness such as stages of sleep. However, interpretation by the human electroencephalographer must always suffer certain inherent difficulties. In this special case he is faced with the formidable task of transforming one or more wiggly inked lines, from recordings lasting from 1 min to 50 hours, into word symbols of a paragraph or a page. The present need is for a more efficient and exact type of transformation of the analog signal into digital rather than word symbols.

The tracing is read in sections recorded at the rate of 1 page or so per 20 sec and may be flipped and scanned in as little as 5 percent of recording time. Processing of the tracing in this way loses much of the information carried by subtle changes; the necessarily qualitative nature of subjective impressions results in several crucial handicaps. It is not practical to compare directly the exquisite details of long records taken continuously on the same subject, so that evaluation of relative change in state of consciousness or stage of sleep is made more difficult. Qualitative data handicap evaluation of a record relative to other subjects or populations as well as make it almost impossible to establish solid statistical levels of confidence for impressions or diagnoses.

For future application in monitoring of space flights, the fact that qualitative data compel the use of an expensive high-level human "computer" for the reading and interpretation of the record is important because the limited availability of time from such trained personnel must always restrict the duration and number of samples that can be interpreted. Continuous, long-term, on-line interpretation is prohibitive.

The inherent inadequacies of subjective data manipulation, and certain shortcomings of several other analytical systems, are overcome by period analysis as employed in analysis of EEG's from space in the following important aspects.

High resolution in time—Period analysis can resolve changes in the signal faster than the EEG pattern can reflect a given state of consciousness. The 10-sec epoch utilized in this study has proved satisfactory for analysis of most experimental sleep patterns; however, abrupt transient changes seen in these records strongly suggest that the analysis should have been in epochs of 1 sec or less.

Long-term continuous analysis—Period analysis can process the signal continually over a time commensurate with the duration of the flight. While it is possible that a particular system of analysis may yield information from relatively few and short samples, continuous analysis is highly desirable.

Multiple channels—Period analysis can interrelate in a meaningful way to two or more simultaneous signals. While single-channel analysis yields important and reliable information, analysis of multiple simultaneous signals significantly increases neurophysiological information and the level of confidence in interpretation.

Relevance to current interpretation of the EEG—While analytical considerations do not require that a system of analysis employ parameters having meaning in the context of EEG interpretation, and while a system may process in such a way that the analyzed data cannot be directly related to current reading, it is preferable for certain of the analyzed parameters to transform directly into signs and interpretations now employed by electroencephalographers. By the same token, the analyzed data and the analytical display should not be more complex nor require greater effort for interpretation than the primary EEG signal itself.

METHOD

Recording of the Analog Data

The recording sites for the EEG were selected and prepared according to reported procedures (ref. 6). The exact positions of the four perforated-electrode sites correspond to the following measurements (ref. 7):

(1) Channel 1—Midline-central site—7.8 in. from the external auditory meatus in the coronal plane and 7.9 in. anterior to the inion in the sagittal plane; midline-occipital site—1.6 in. superior to the inion in the midsagittal plane

(2) Channel 2—Left-central site—3.1 in. to the left of the midline-central site in the coronal plane; left-occipital site—1.4 in. to the left of the midline-occipital site

The master analog magnetic tapes were recorded on special equipment* at 0.0293 in./sec and rerecorded through four steps in a format compatible with a Precision Instrument, 14-channel, record-playback system. The tapes were replayed at $1\frac{7}{8}$ in./sec with output simultaneously recorded on an eight-channel Grass model-III electroencephalograph and fed to the analog-to-pulse-width converter for period-analytic processing.

Period-Analytic Processing

Period analysis of the EEG as a data-reduction process has been reported (refs. 8 and 9), but one must understand that period analysis yields these three basic parameters: major period, intermediate period, and minor period. These three periods respectively code the base-line cross of the primary EEG or dominant activity, the peak-and-valley activity, and the very low-amplitude inflection points of the EEG signal. They are subsequently distributed over three spectra of "equivalent" frequencies in 10 bands per spectrum (table 1).

The most elegant statistics of period analysis are the counts per second of the major, intermediate, and minor periods. This extremely economical process, ideal for on-line data reduction in real time with minimum instrumentation, offers a reliable index of state of consciousness and stage of sleep. A further smoothing step, generating a single statistic that is generally monotonically related to arousal, is the simple process of summing all the counts for what we refer to as the total count. As in all smoothing

*Cook Electric Company, Morton Grove, Illinois.

TABLE 1.—*Equivalent Frequencies in Hertz*

Band	Period		
	Major	Intermediate	Minor
1	1–4	1–4	1–10
2	4–6	4–6	10–20
3	6–8	6–8	20–30
4	8–10	8–10	30–40
5	10–12	10–12	40–50
6	12–18	12–18	50–60
7	18–24	18–24	60–70
8	24–35	24–35	70–80
9	35–50	35–50	80–90
10	50–100	50–100	90–100

procedures some information is lost in the total-count statistic, and in some cases of ambiguous interpretation one is forced to the independent counts for clarification. In studies involving sleep it is convenient to weight the major-period count by a factor of four, the intermediate-period count by a factor of two, and the minor-period count by a factor of one in order to accentuate the monotonic relation.

In the 10-sec-epoch smoothing mode, period analysis yields 33 parameters which statistically characterize each 10-sec sample of the EEG record. Each parameter is read as a two-decimal digit value, with the 30 band parameters of frequency distributions expressed as percentage of time occupied by activity in that band, and the three counts expressed as absolute counts per second for each of the three periods. The 33 two-decimal digital statistics characterizing each 10-sec epoch and time identification are logged on incremental magnetic tape at 200 bits per inch in a format compatible with general-purpose digital computation on the IBM-7094. The flow diagram for the data-analysis and data-logging system used is schematized in figure 1.

While state of consciousness should be interpreted from the period-analytic parameters at the time of on-line analysis, more complicated classification procedures were investigated in an attempt to improve the interpretation. Unfortunately these procedures require programming of a general-purpose digital computer for identification of particular classes of electroencephalographic records in order to automate interpretation; they attempt to define mathematically, still by way of period-analytic characterization, certain neurophysiological states as expressed by a set of EEG patterns. This general problem of pattern recognition is becoming increasingly important in psychophysiological and behavioral studies because of the power of these methods. However, such manipulations suffer the disadvantage that the complex relational patterns, that often serve as the criteria for recognition, may be so highly abstract that interpretation in

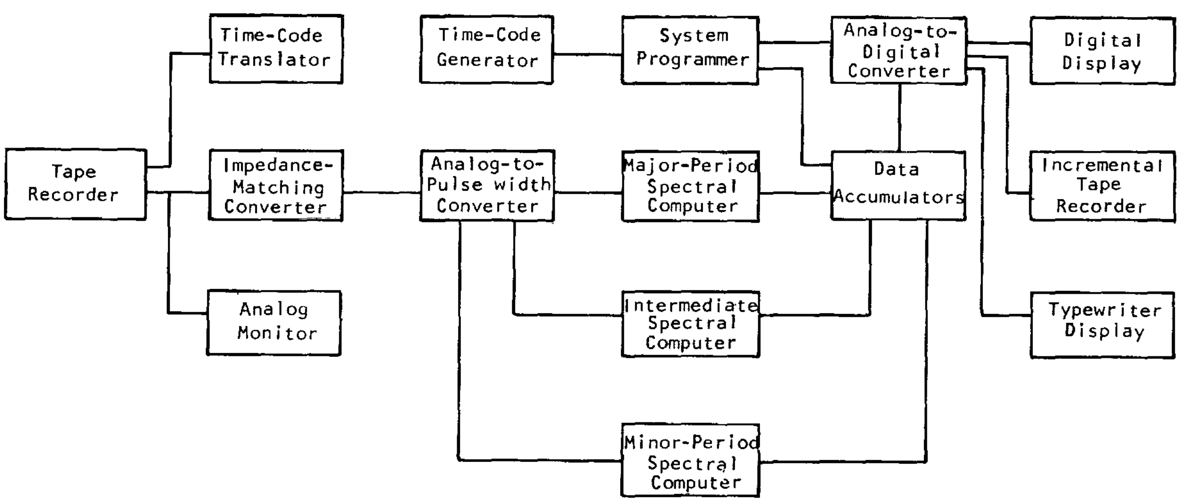


FIGURE 1.—Flow diagram for the data-analysis and data-logging system.

behavioral or neurophysiological terms is almost impossible.

Discriminant Analysis

The technique of pattern recognition here utilized for classification of states of consciousness is discriminant analysis (ref. 10). In discriminant analysis, as in most pattern-recognition programs, criterion groups or training sets must first be established by some selection procedure that is independent of the pattern-recognition logic itself. Thirty-five 10-sec epochs for each of 16 different states were selected by clinical interpretation of the analog record. Figure 2 presents representative 10-sec analog records of three of these different states of consciousness. The initial set of numbers at the lower right of each sample gives the total counts of the major, intermediate, and minor periods; for instance, 11:24:37 in the first record (fig. 2) indicates a count of 11/sec in the major period, 24/sec in the intermediate period, and 37/sec in the minor period. The other numbers below each record convey flight time, a state of consciousness, and the probability associated with this record being a member of the particular class identified. For example, in the first recording of figure 2 the flight time is 030851; the state identified as

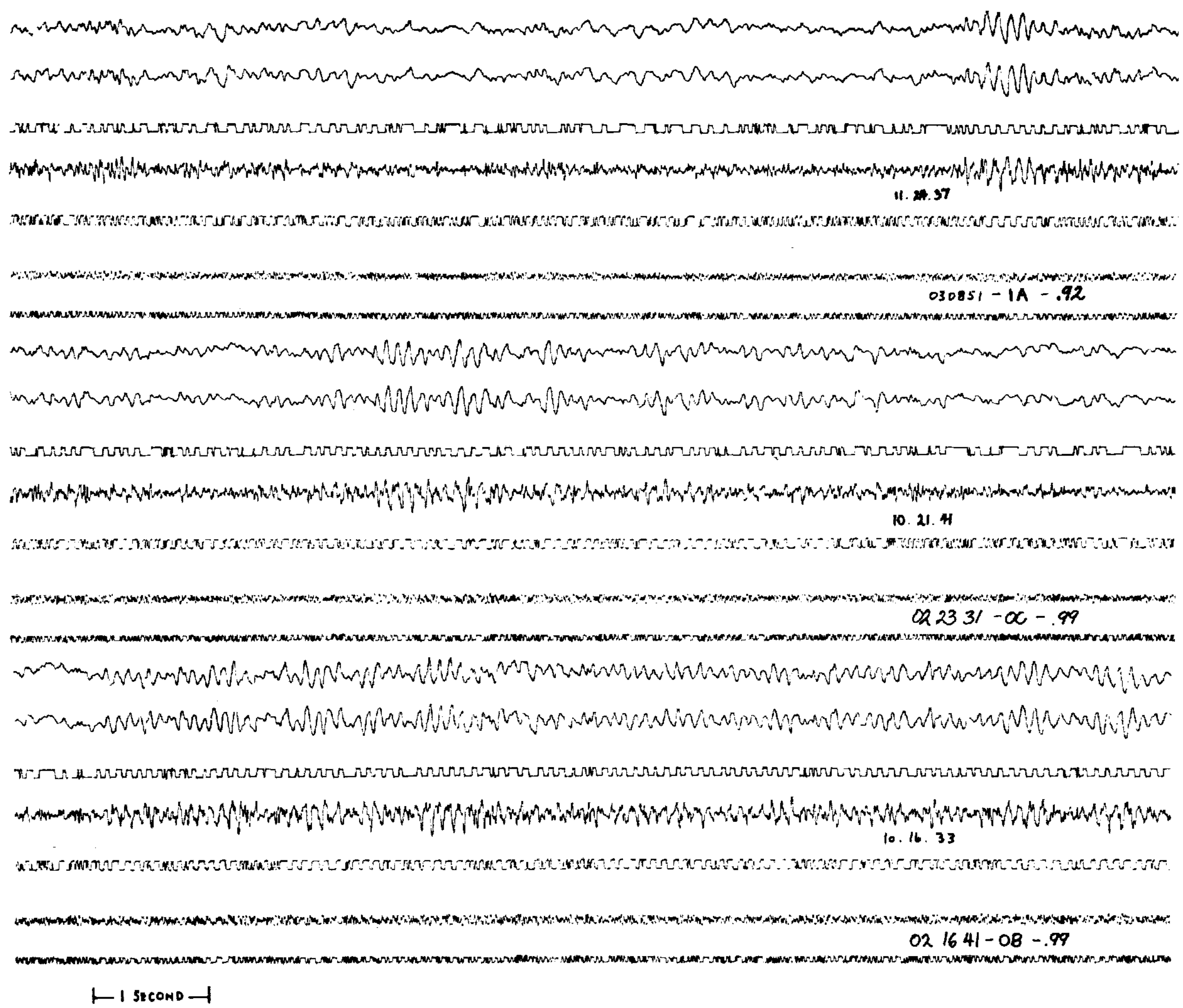


FIGURE 2.—Analog EEG for three operational states of consciousness: early stage-I sleep (1A), poorly organized alpha (0C), and well-organized alpha (0B).

"1A" is stage-I sleep; and the probability that this 10-sec epoch belongs to class stage-I sleep is 0.92, as determined by the discriminant analysis.

The classes established by clinical interpretation of the EEG were grouped into 16 operational categories which represent five states of consciousness. We cannot emphasize too strongly that these classifications are based on the EEG signal only since there was no independent measure of such states as "eyes open" or REM. Table 2 relates the classifications of this report to the standard Dement-Kleitman classifications. For adequate interpretation of the sleep data, results of a comparative study with a simulated Apollo flight are presented.

RESULTS

The fact that the period counts, weighted total counts, and discriminant-function classifications of the 50 hours of EEG have definitive relations to the events occurring during the mission shows that these parameters are sensitive indicators of the state of consciousness during space flight. Additionally these results demonstrate the qualitative and quantitative changes in the sleep patterns during flight. Of all the operations performed on the data, the weighted total count gives the most economical and reliable indicator of the state of consciousness.

The relations of the mission's events to the discriminant-function analysis and to weighted total counts of the EEG will be used to demonstrate the interpretive process. Where these two characterizations of the state of the command pilot are not congruent, more-specific data from other parameters of the analysis will be presented to clarify the interpretation. In order to follow the time-line profile one must interpret the principal classifications of the different electroencephalographic states as behavioral states.

TABLE 2.—*Equivalent Classifications*

Discriminant	Classical (activity)	
	Predominant	Secondary
1A	I-Submergent	
2A	I-Submergent with alpha	
2B	I-REM	
2C	II	
3A	II	
3B	II	III
3C	III	
4A	III	IV
4B	IV	III
4C	IV	

Principal Classifications of the Behavioral States

No classification program can be better in terms of sensitivity and selectivity of discrimination than the basic parameters or descriptors that characterize the data. Therefore, it is of great importance to appreciate the power of the 30 period-analytic bands in characterizing each category.

Table 3 lists the mean percentage of time in each band for each of the 16 categories; the total counts are included but were not used as descriptors in the discriminant analysis. The categories are designated by the operational definition used for selection of the criterion samples and by the Dement-Kleitman equivalents. On the basis of the detailed findings presented in table 3, the interpretative keys for the principal classifications are now outlined.

Artifact—Heavy-muscle artifact shifts the frequency histograms of all three periods to the right, with accentuation of all high-frequency components. Moderate-muscle artifact tends to be contaminated with slow-wave-movement artifact which shifts the major-period histogram to the left; the high-frequency muscle components continue to shift the intermediate and minor periods to the right even when the muscle artifact is "moderate."

Arousal—The T_1-category may be interpreted as a state of nonspecific neurophysiological arousal in contrast with the relatively specific visual arousal of the eyes-open category. Both arousal states show increase in slow components, but delta tends to predominate in the T_1-category while slow theta is dominant in eyes-open category. The eyes-open state shows twice as much 24-to-35-Hz activity as does T_1; this accentuation of the relatively high-frequency beta component may characterize the specific visual arousal of

TABLE 3.—*Percentages of Time for 10 Bands Per Period and Counts Per Period*

Period	Percentage time per band										Counts*	
	1	2	3	4	5	6	7	8	9	10	Period	Weighted
					MB—Heavy muscle artifact							
Major	13	13	11	6	7	15	13	15	8	7	17	68
Intermediate	1	0	0	1	2	13	16	40	23	24	39	78
Minor	1	5	19	27	11	11	15	13	7	2	50	50
											(106)	(196)
					MA—Moderate muscle artifact							
Major	28	25	14	8	5	10	8	6	2	2	10	40
Intermediate	1	1	4	4	6	26	21	29	13	15	30	60
Minor	1	12	24	28	8	6	10	11	7	2	47	47
											(87)	(147)
					T1—10 min before to 15 min after lift-off							
Major	32	27	16	8	5	10	7	4	2	4	11	44
Intermediate	1	2	6	7	7	32	21	27	10	13	29	58
Minor	1	16	28	29	7	6	9	10	9	4	45	45
											(85)	(147)
					EO—Eyes open							
Major	16	19	14	9	8	14	13	8	4	3	13	52
Intermediate	1	0	1	2	3	22	20	35	16	16	32	64
Minor	0	9	26	32	8	6	9	10	6	2	46	46
											(91)	(162)
					OB—Moderately well-organized alpha							
Major	2	4	9	25	29	29	7	2	1	0	11	44
Intermediate	1	0	2	17	24	41	12	11	4	4	18	36
Minor	3	34	27	23	5	3	4	5	4	1	36	36
											(65)	(116)
					OC—Poorly organized alpha							
Major	5	7	15	20	21	24	9	3	1	0	11	44
Intermediate	1	0	4	13	17	42	16	15	5	5	19	38
Minor	2	31	29	26	5	3	4	5	4	2	37	37
											(67)	(119)
					1A—Submergent stage-I							
Major	9	14	15	16	12	20	12	5	2	1	11	44
Intermediate	1	0	2	8	11	35	21	25	8	7	23	46
Minor	1	22	29	29	6	3	5	6	4	1	39	39
											(73)	(129)
					2A—Submergent stage-I with alpha							
Major	7	15	16	16	14	19	9	4	2	0	11	44
Intermediate	0	0	3	9	12	36	19	22	7	7	22	44
Minor	1	24	31	28	5	3	4	6	4	2	38	38
											(71)	(126)
					2B—Stage-I REM							
Major	19		19	10	7	11	10	4	2	1	10	40
Intermediate	1		4	8	9	29	23	27	8	7	24	48
Minor	0	18	33	32	6	3	4	5	5	2	39	39
											(73)	(127)
					2C—Stage-II (light)							
Major	29	28	19	8	5	7	6	3	2	0	8	32
Intermediate	1	2	7	10	11	30	19	22	8	7	23	46
Minor	1	21	30	31	6	3	5	6	5	2	40	40
											(71)	(118)
					3A—Stage-II (moderate)							
Major	35	30	14	8	5	8	5	2	2	0	7	28
Intermediate	1	2	9	12	12	32	17	19	7	7	22	44
Minor	1	22	29	30	6	3	5	6	5	2	41	41
											(70)	(113)

TABLE 3.—*Percentages of Time for 10 Bands Per Period and Counts Per Period*—(Concluded)

Period	Percentage time per band										Counts*	
	1	2	3	4	5	6	7	8	9	10	Period	Weighted
					3B—Stage-III (light); stage-II (deep)							
Major	40	35	14	7	4	7	4	2	2	0	7	28
Intermediate	1	5	11	15	13	30	13	16	6	6	20	40
Minor	1	24	29	28	5	3	4	6	5	2	39	39
											(66)	(107)
					3C—Stage-III (moderate)							
Major	45	29	12	6	5	8	5	2	1	0	7	28
Intermediate	2	5	11	14	14	30	14	14	6	6	20	40
Minor	2	27	26	26	5	3	5	6	5	3	39	39
											(66)	(107)
					4A—Stage-IV (light); stage-III (deep)							
Major	54	26	10	6	4	6	3	1	1	0	5	20
Intermediate	2	9	18	19	15	26	10	11	4	5	17	34
Minor	3	30	27	25	5	3	4	6	6	3	37	37
											(59)	(91)
					4B—Stage-IV (moderate)							
Major	60	21	10	6	4	4	3	1	1	0	5	20
Intermediate	2	11	21	22	15	24	9	10	3	4	16	32
Minor	3	33	24	25	5	3	3	6	6	3	37	37
											(58)	(89)
					4C—Stage-IV (deep)							
Major	72	15	6	3	2	3	2	1	0	0	4	16
Intermediate	6	16	22	20	13	21	8	9	3	4	16	32
Minor	4	33	25	25	4	3	4	6	6	3	38	38
											(58)	(88)

*Total counts for three periods appear in parentheses.

the eyes-open situation. The predominant 30-to-40-Hz activity in the minor period further defines this specific visual component as being quite well organized and primarily in the fast-beta range. Compared to the beta component of specific arousal, the T_1-state tends to show increased superimposed activity in the 12-to-18-Hz range, which is demonstrated to be a well-organized wave shape by the comparatively high 10-to-20-Hz component of the minor period.

Resting—The eyes-closed (presumably resting) categories show the expected high percentages of alpha activity with very little in the way of slow components. The resting state shows an unexpectedly high percentage of 12-to-18-Hz activity so well organized that it remains modal for the superimposed wave shapes. The very-high-frequency components of 50-to-80 Hz have diminished in this eyes-closed state by 50 percent or more when compared to the arousal categories.

Sleep states—Progressively deeper sleep is characterized by a rather well behaved increase in slow components at the expense of higher frequencies. The progressive slowing first appears in the dominant primary activity and then in the superimposed activity of the intermediate period; finally, as the slow-wave components become synchronous, they are reflected in the minor period. The mean value of 4 percent in band-1 (1–10 Hz) of the minor period, seen in deep stage-IV sleep (4C), is quite significant and probably pathognomonic of a type of neurophysiological activity that can exist only in this very special state of extremely slow-wave sleep. The depth of sleep can be directly related to the shift to the left of the three period histograms.

An exception to this rule occurs in light sleep, particularly in that presumed REM state identified as 2B. In this stage the orderly progression of the shift to the left has been interrupted by a relative shift to the right and an increase in relatively high-frequency activity as compared to the previous sleep stages. Of particular interest is the extremely high activity in the 20-to-30- and 30-to-40-Hz ranges in the minor period. This component may represent a specific activation state characteristic of this rather unique stage of sleep associated with rapid eye movements.

Time-Line Profile

We may now turn to the time-line profile generated from the midline EEG of the command pilot and recorded from some 10 min prior to lift-off through the first 54 hours of the flight.

Figure 3 shows the discriminant-analysis time-line profile, plotted according to the state of consciousness in 5-min data points, for the first 28 hours of flight. The state-of-consciousness characterization of the 5-min sample was determined by the most frequently occurring, or modal, 10-sec epoch; muscle artifact is plotted only if it is the exclusive characterization in all 10-sec epochs of the 5-min sample. In majority-vote logic of this sort, provision must be made for a tie vote if two or more categories occur with the same frequency; in the case of a tie, the category representing the lower state of consciousness is taken to characterize the 5-min sample.

Figure 4 illustrates the weighted total-count profile plotted as 5-min sample points against the first 28 flight hours. Only time points of particular interest will be discussed, but the interpretation of this index as monotonically increasing with increased arousal and monotonically decreasing with depression in state of consciousness is so straightforward that the reader is invited to render his own interpretation for any given point in time. A paradoxical area of interpretation is found in stage-I ("paradoxical") sleep (ref. 11).

The time-line profile of selected individual bands will be used occasionally to illustrate certain points in particular time samples; at points of ambiguous interpretation the independent major-, intermediate-, and minor-period counts must be called on for clarification. Light-dark cycles and orbital revolutions are indicated (circled dots) on all time-profile graphs.

From $T-10$ to $T-5$ Minutes

Mission events—A behavioral state of alertness would be expected to accompany rather intensive preparations for lift-off. It is known that there is some normal increase in tension at lift-off (ref. 12). The EEG recording begins approximately 10 min before lift-off ($T-10$ min).

Discriminant analysis—Since the training set of the T_1-category was drawn primarily from this time, it is not surprising that the modal profile finds such T_1-epochs to be the most numerous category during these 5 min.

Weighted total counts—A relatively high level of arousal is indicated by the total-count reading of 162.

Band parameters—The time-line profile of delta activity is illustrated by the upper portion of figure 5. The reading in this 5-min period is relatively high at 27 percent; this is the component that previously has been interpreted as a special case of arousal. The lower portion of figure 5 shows the minor-period activity from 10-to-20 Hz also to be relatively high in this state of consciousness.

Mission events—Anticipation of imminent lift-off and decrease in the rate of preparatory activity may have produced a special state of vigilance with a strong behavioral component of inhibition.

Discriminant analysis—The discriminant classification continues to show the T_1-states although there has been considerable change in the total counts. The stability of the modal smoothing procedure, with some loss of sensitivity, is clearly seen in the contrast of these two indices.

Weighted total counts—The difference in the level of activation, between the 5 min immediately prior to lift-off and the 10 min before lift-off, is seen as the total counts drop from 162 to 144. The mean value of the two samples preceding lift-off, 152, is used in this report to establish the midpoint of what we shall refer to as the T_1-zone which ranges from 150 to 154 counts. This T_1-zone is interpreted as an arousal state reflecting special vigilance; in the weighted total-count graphs it is the lightly shaded bar from 150 to 154.

Lift-off to $T+5$ Minutes

Mission events—The acceleration profile of this mission shows its first peak of some 5.5 g approximately 2.66 min after lift-off.

Discriminant analysis—The state of consciousness now changes to an eyes-open classification.

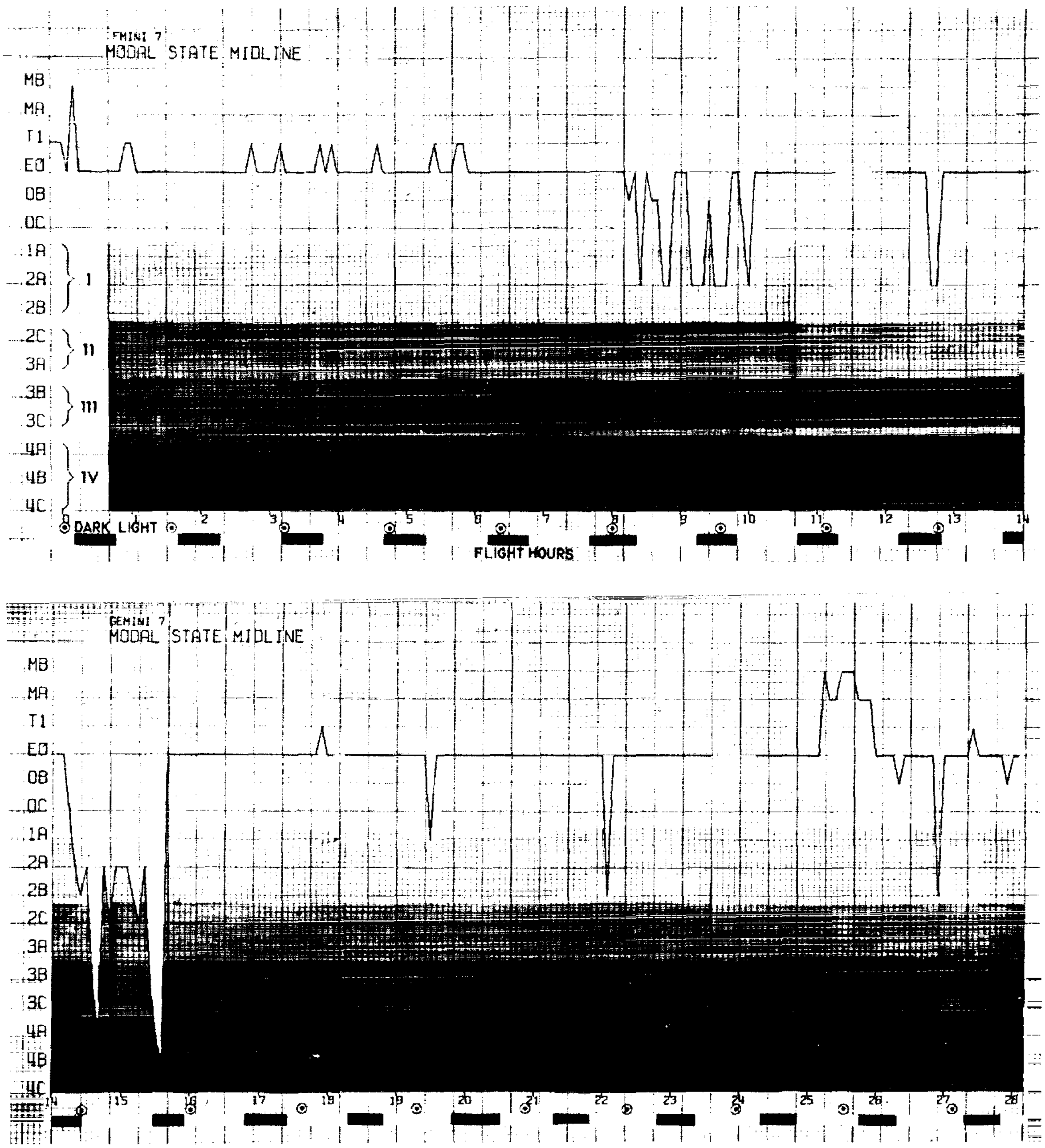


FIGURE 3.—The modal state of consciousness for every 5 min for the first 28 hours of the flight.

Weighted total counts—A count of 163 does not indicate that the event of lift-off itself has markedly changed the level of arousal.

From $T+5$ to $T+10$ Minutes

Mission events—The acceleration curve shows its second and highest peak, about 7.33 g, about 5.66 min after liftoff. The astronauts report that this period of second-stage-engine cutoff is a "crisp event" when the g-level suddenly drops to zero (ref. 13).

Discriminant analysis—The extremely rare exclusive heavy-muscle state appears in response to maximum g-forces.

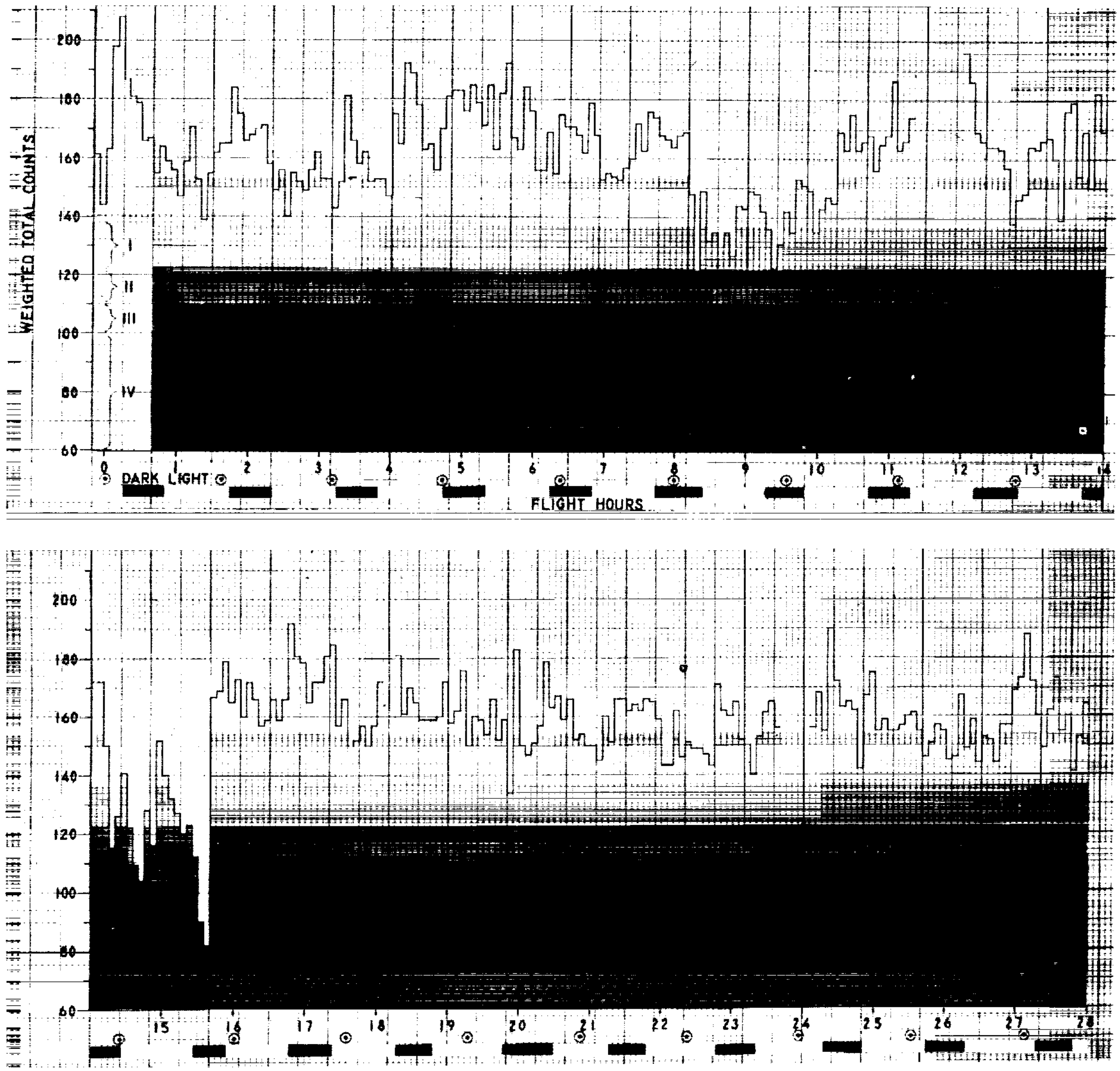


FIGURE 4.—Weighted total-count graph for the first 28 hours of flight.

Weighted total counts—The striking increase of the total counts to 198 reflects the effect of maximum *g*-forces.

From $T+10$ to $T+60$ Minutes (00:10 to 01:00)

Mission events—It would of course be of extreme interest to know in exact detail what transpired to force a return to the level of vigilance represented by the T_1-state. The astronauts consider critical the early phase of the flight when they are adjusting to their new environment (ref. 13). The rate of verbalization by the command pilot in communication with the ground is maximal during this early part of the mission; there is relative increase in rate of transmission during the T_1-state.

During all but 16 hours of the mission the oxygen-to-water differential-pressure warning light of section-2 indicated a beyond-limits oxygen-to-water pressure across the water sepa-

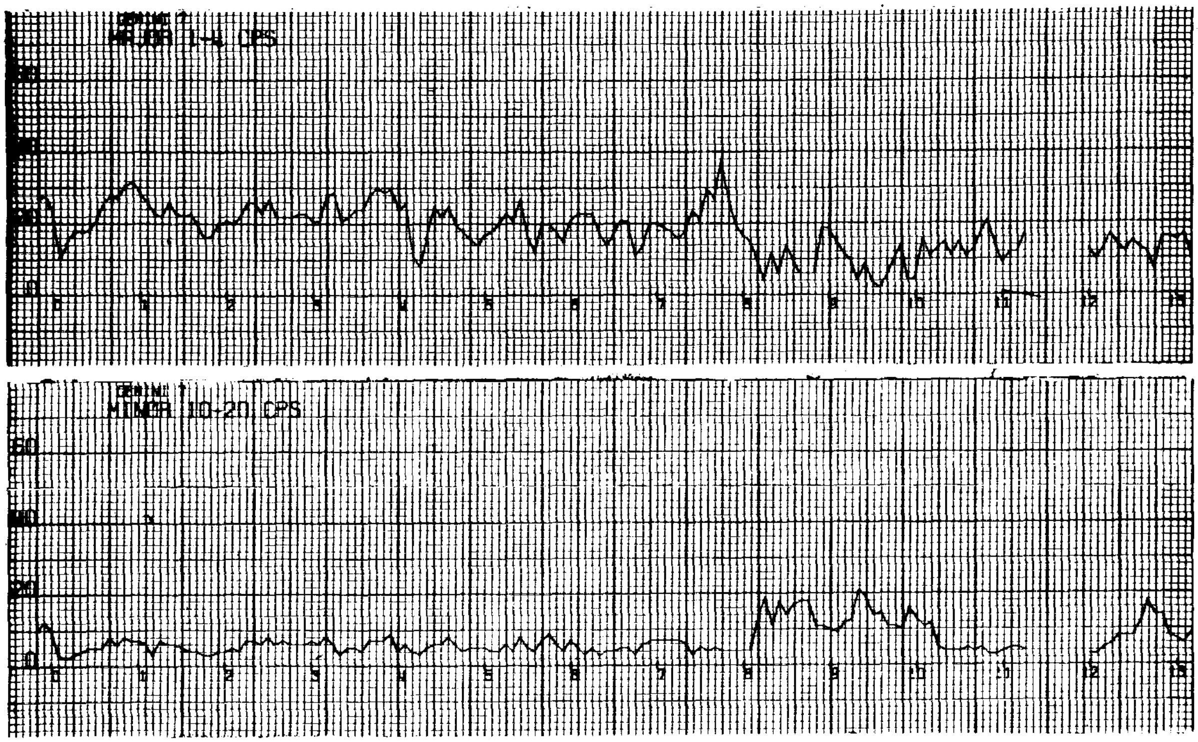


FIGURE 5.—Comparison of the major-period delta-index band-1 (1 to 4 Hz) with the minor-period band-2 (10 to 20 Hz) to illustrate the rather unusual slow activity occurring 10 min before lift-off.

rators (ref. 14). Considering the important part that a fuel-cell beyond-limits warning light had played in Gemini V (ref. 15), it must be assumed that the spurious warning light in the early part of this later flight was a signal of some consequence to the command pilot. We may speculate that the differential-pressure warning light was a highly alerting signal in the behavioral sense.

Discriminant analysis—Most of the 1st hour is occupied by the eyes-open modal state, but just before the end of this period there is a return to the T_1-state for 10 min.

Weighted total counts—The peak arousal of 208 counts occurred when the astronaut was experiencing his first continuous zero-*g*; it may represent one phase of accommodation to this novel environment. The relatively low count value of 154 between 00:45 and 00:50 occurred when the extremely high rate of verbal communication from the command pilot to the ground diminished almost to silence. A higher rate of verbalization begins during the subsequent 5 min, with a concurrent jump to 164 counts. Between 00:55 and 01:00 the reading of 147 counts is very close to the T_1-zone, and it is at approximately this time that the discriminant analysis also indicated a T_1-state.

From 01:00 to 02:00

Mission events—The T_1-state of this hour again coincides with relatively increased verbalization from the command pilot to ground. In the light-dark cycle, daylight had begun about 10 min before 01:00 and ended about 15 min before 02:00. During the first two dark cycles the downward swing of the EEG cycle seems to coincide with the onset of darkness.

Discriminant analysis—This hour is almost exclusively characterized by the eyes-open state, except for the one 5-min sample between 01:15 and 01:20 which again shows the T_1-state.

Weighted total counts—The weighted-count level of 139 shows a remarkable degree of relaxation at 01:25, but the level of arousal progressively increases during the next 25 min to peak at a value of 184 before cycling down to 140 at 02:35.

In behavioral terms, the weighted total-count level of 140 represents relative inhibition; in terms of performance such relative inhibition is a "good" or "bad" state depending on whether the stimulus field of the moment calls for considered action or relaxed inaction.

From 02:00 to 06:00

Mission events—The second and third orbits are seen to have been completed at 03:15 and 04:45. It is reasonable to suppose that any change in the state of arousal, associated with the orbital schedule, should now begin to be evident. At some time during the mission the windows were screened with foil from the food packs, in addition to the Polaroid screen, to keep the cabin comfortable during sleep periods (ref. 16). Such screening would of course reduce whatever EEG-arousal effects the light-dark cycle might engender.

Discriminant analysis—During this 4-hour period the eyes-open state predominates, but the T_1-state appears seven different times for a total T_1-time of 40 min.

Weighted total counts—A phase-locked relation appears to be developing between the weighted total-count profile and orbital revolutions, suggesting that an arousal cycle is initiated at the beginning of each new revolution and increases for about 15 min to a peak value before relative relaxation as the revolution is completed. The weighted total counts fluctuate within the T_1-zone for 35 min of this time period. The fact that the weighted total counts do not identify exactly the same samples as those already identified by the discriminant analysis as T_1 shows that the indices are employing slightly different criteria for recognition.

From 06:00 to 12:00

At the onset of sleep, slightly after 08:00, the command pilot has been awake for about 15 hours; this time also corresponds to his "biological time" of about 2300 hours. Both of the foregoing factors recommend interpretation of this 2-hour sleep cycle as being tantamount to the first 2 hours of a regular night's sleep for him.

The flight plan was designed to allow the crew to sleep during hours generally corresponding to night at Cape Kennedy. This plan was followed since the sleep program on previous flights had not worked well because of flight-plan activities and the fact that the crew tended to retain their Cape Kennedy work-rest cycles, with both crewmen falling asleep during the midnight-to-0600 period of night at Cape Kennedy (ref. 16).

Discriminant analysis—The T_1-state does not appear in this 6-hour portion of the mission. The eyes-open state occupies the first several hours, and beginning at 08:15 we see the first indication of relaxation with the appearance of well-organized alpha suggesting that the eyes are closed. From 08:20 to 08:25 the first intimation of light sleep appears, immediately followed by an eyes-open and then by a resting eyes-closed state. Ten minutes of light sleep is noted between about 08:40 and 08:50. During this first sleep cycle between 08:00 and 10:00 we find five episodes of this very minimal depression in the state of consciousness with a total light-sleep time of 50 min.

Weighted total counts—The onset of light sleep is seen at 08:20 as a weighted count value of 122, indicating that stage-II sleep has occurred. Another episode of stage-II sleep apparently occurs at 09:25, but caution should be exercised in differentiating stage-I from stage-II sleep on the basis of weighted total counts only; this is one of the paradoxical exceptions to the monotonic relations between total counts and level of arousal. A mixture of hypersynchronous alpha, associated with the lightest stage-I sleep (ref. 17), may have weighted total counts ranging in the 80's; low-voltage fast activity, also found in stage-I sleep, may have weighted total counts ranging around 127. Such a mixture yields a combined total count that can incorrectly indicate sleep levels deeper than stage-I. The possibility of such an indication disappears as higher and higher time-resolution is employed in analysis of the EEG; as time epochs of 1 sec and less are employed, the EEG tends to become a "pure culture" representation of the momentary state of consciousness.

Here is an example of an ambiguous interpretation in the smoothed index of weighted total counts that must be resolved by reference to the unsmoothed, independent, major-, intermediate-, and minor-period count profiles. Figures 6 and 7 present the count profiles for the three periods plotted as average counts per second in 5-min

FIGURE 6.—The period-total counts for the first 28 hours of flight, displayed with the average count for every 5 min.

data points for the full 54 hours of flight. The major-period count was 9/sec while the intermediate and minor counts were 22 and 42, respectively. These values indicate that the distributions of the independent counts are incompatible with stage-II as indicated by the weighted total counts, and corroborate the interpretation that this activity was compounded stage-I.

From 12:00 to 18:00

Mission events—No direct information is available concerning mission events during this time, but the recurrence of the T_1-state prior to 18:00 leads to speculation that some event was again provoking a state of special vigilance.

Discriminant analysis—A 10-min episode of stage-I sleep is seen beginning at 12:40, followed by a return to the eyes-open state until the beginning of a new sleep cycle at 14:15. At 14:25 the deeper stage-I sleep of category-2B, the presumed REM and dreaming state, is reached for the first time. Stage-2B is followed after about 10 min by stage-3B, the first occurrence of stage-III sleep. At 14:45 the first stage-IV sleep is noted with the appearance of category-4A. After only 5 min in stage-IV sleep, the state of consciousness cycles almost immediately to the

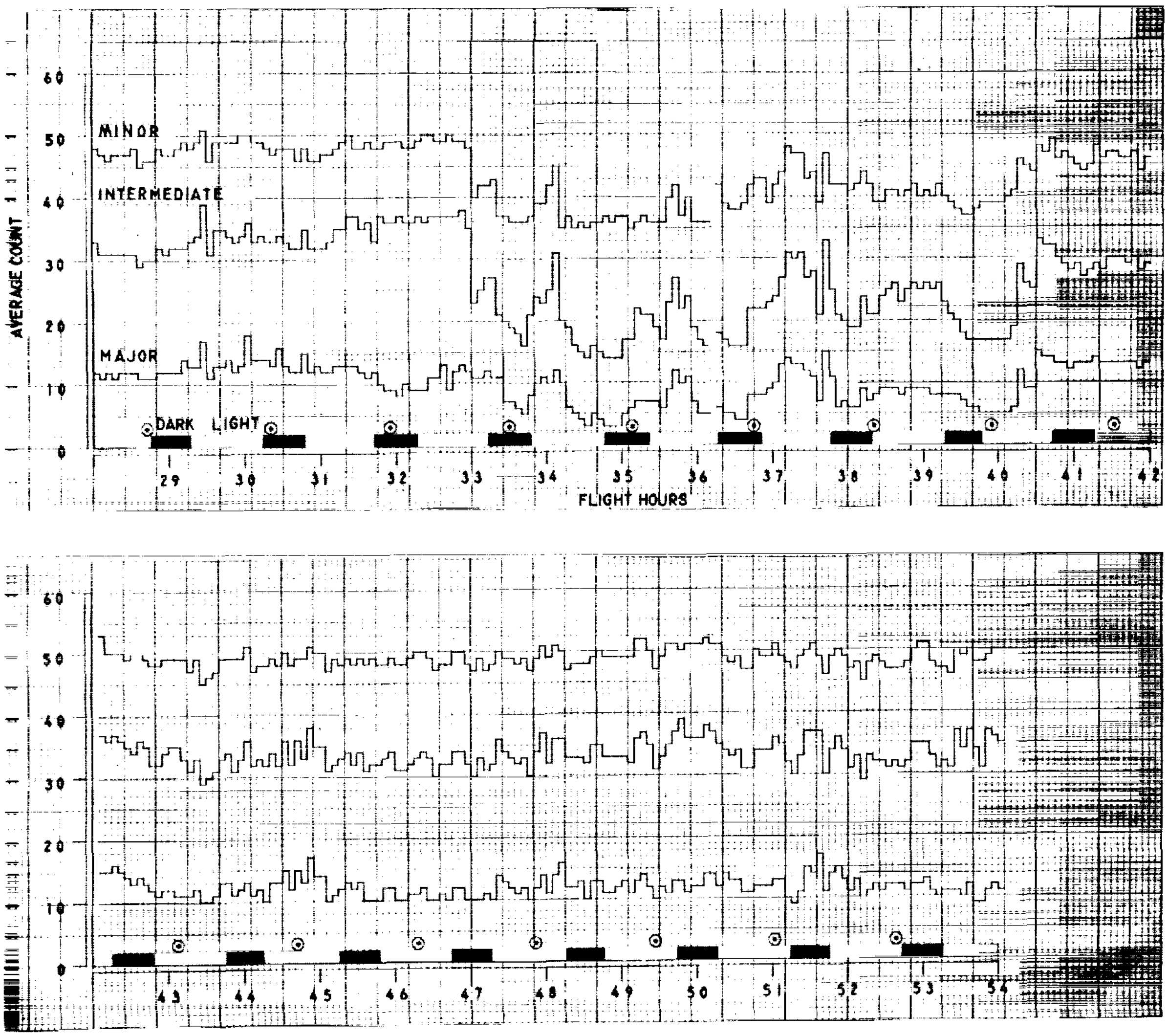


FIGURE 7.—Extension of figure 6 between 28:00 and 54:00.

lighter sleep of stage-I and early stage-II, which continues for some 35 min. Moderately deep stage-IV sleep of category-4B lasts for only 10 min at 15:35, and in the next 5 min the state of consciousness returns to the eyes-open category. For 5 min, just prior to 18:00, again we see the now-rare T_1-state.

Weighted total counts—The weighted total-count profile comes suspiciously close to the stage-I sleep level during 12:45, and the discriminant-analysis profile also found stage-I sleep to be occurring. Forty minutes later another depression in the state of consciousness again comes quite close to stage-I sleep; this depression had not been detected by the discriminant analysis.

This now-familiar sleep segment accords with previous interpretations except for the two episodes of waking for 5 min at 14:30 and for 10 min between 14:55 and 15:05, the discriminant analysis had classified these samples as stage-I sleep. The value at 14:30 is an example of a corollary to the paradoxical exception in the interpretation of the weighted total count that was discussed earlier. Relatively "pure culture," low-voltage, fast activity of stage-I sleep, particularly of stage-I REM, may yield a relatively high weighted total count that may incorrectly indicate a waking state. In such circumstances, reference to the independent period counts is necessary for clarification of the interpretation.

The clear indication of waking with a weighted total-count value of 152 at 14:55 carries period counts of 13:28:44 which corroborate the fact that this is an awake level. This second discrepancy, between the discriminant-analysis classification and the weighted total-count assignment of state of consciousness, may be used to illustrate a point of considerable practical importance. This discrepancy can also be resolved by reference to the discriminant-analysis profile of the second channel of EEG recorded during this flight.

The discriminant-analysis classification, of the left parieto-occipital lead, agrees with the weighted total count of the midline lead that this is an eyes-open waking state. This fact demonstrates the practicality and even necessity of multiple electroencephalographic leads; information from the second lead is not redundant if any ambiguity is present in interpretation of the single-channel signal. This argument holds quite aside from additional neurophysiological information that one may expect from multiple-site recordings and interrelations.

From 18:00 to 24:00

Mission events—The summary flight plan (ref. 18) indicates that the fuel cell was purged at 20:50. From 21:00 to 21:30 the command pilot was engaged in a vision test, and at 22:25 a radar-transponder test was scheduled, followed by an exercise period.

Discriminant analysis—The eyes-open state occupies the entire 6-hour period except for two episodes of very light sleep appearing at 19:25 and from 21:55 to 22:00.

Weighted total counts—The weighted total counts show that the special state of vigilance, referred to as the T_1-zone, accompanies the above mission events except for the period of exercise. It seems most reasonable to assume that an event, such as purging of the fuel cell, would tend to reactivate the state of arousal that earlier was associated with the differential-pressure warning light; again the important part played by the fuel cell in the flight of Gemini V is recalled. Shortly before 20:00 a count of 134 indicates a period of light sleep. Weak cycling continues in relation to revolutions.

From 24:00 to 30:00

Mission events—The activity profile indicates that from just before 25:00 through 26:00 was a mealtime, and the chewing artifact unquestionably accounts for the EEG findings of 25:00. According to the flight plan a fuel cell was purged at 28:40, and window measurements were taken at 29:40.

Discriminant analysis—A rather unusual range of states is seen during these 6 hours (fig. 3). Extreme-muscle artifact preempts the record for some 45 min from approximately 25:00. Light sleep of the 2B-category is noted briefly at 26:40. Two episodes of the T_1-state occur at 27:10 and 28:35.

Weighted total counts—The extreme-muscle activity from the 25th to the 26th hour is not clearly demonstrable in the weighted total-count profile of figure 4. Again this is an example of a smoothing statistic partially obscuring the information of the basic data. Figure 6 strikingly shows the muscle activity seen in the intermediate and minor periods as extremely high counts, and a slow-wave "chewing" artifact component seen as the concurrent depression of the major-period counts. This "mirror image" of the major-period counts compared to the intermediate- and minor-period counts is a characteristic artifact pattern. It is now of increasing interest that the mission events of this period relate strongly to the state of consciousness reflected by the weighted total count in the T_1-zone.

From 30:00 to 36:00

Mission events—The flight plan indicates that the radar transponder was turned "on" at approximately 31:55 and turned "off" at around 32:10; there was another purge of the fuel cell at 32:10. The summary flight plan schedules the beginning of the command pilot's sleep period at 33:20.

Discriminant analysis—Muscle artifact, alternating with T_1, occurs for some 25 min before 32:00 (see fig. 8, a continuation of fig. 3). Early relaxation of the eyes-closed state begins around 33:00 to develop into very light sleep at 33:10; a brief oscillation back to the eyes-closed state

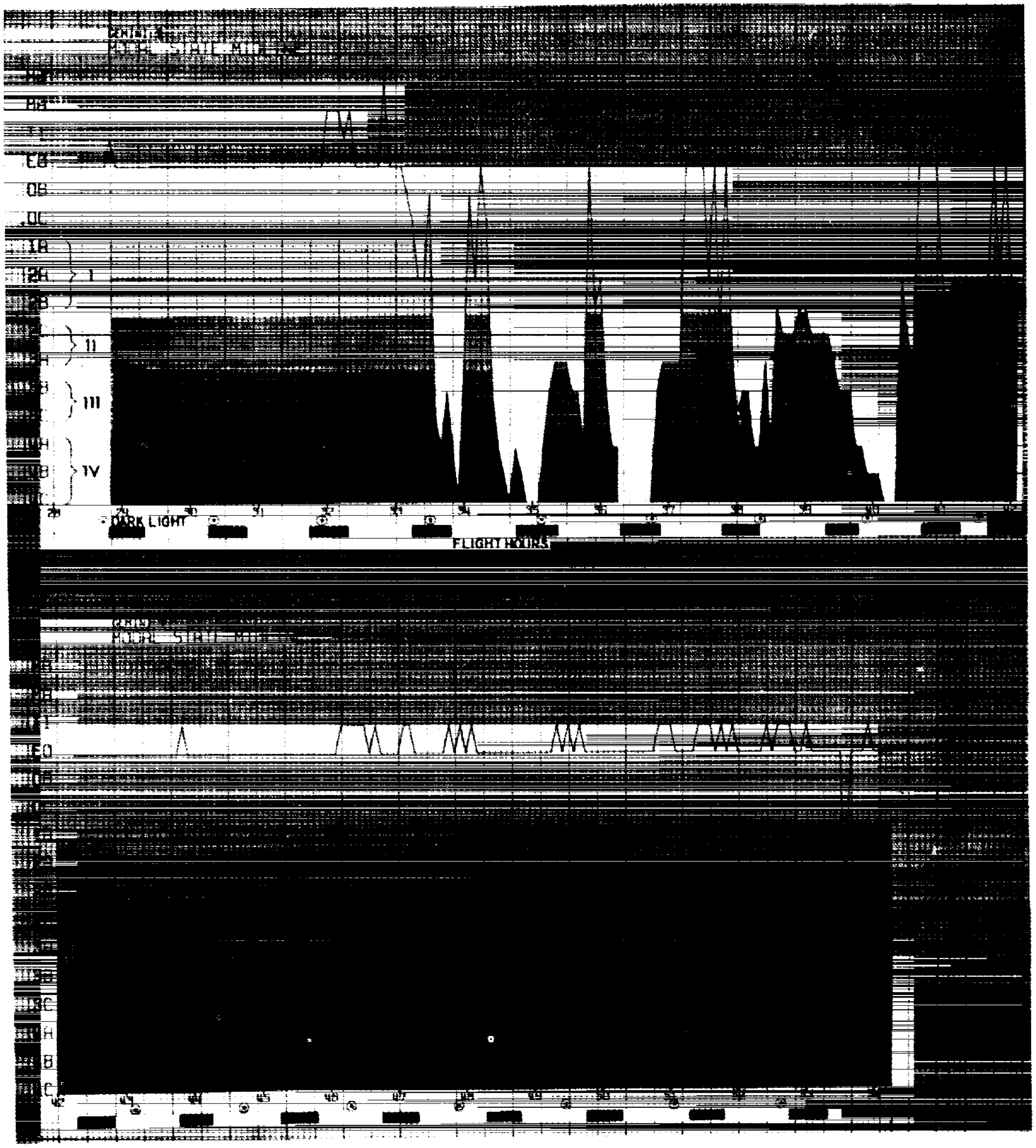


FIGURE 8.—Continuation of figure 3 between 28:00 and 54:00.

is followed at 33:25 by the beginning of the deepest sleep cycle yet seen, dropping to the 4C category of deep stage-IV sleep at 33:35. However, this deep sleep is very brief and is almost immediately interrupted by 25 min of eyes-closed resting and very light sleep. The first deep-sleep cycle of reasonable duration may be said to begin at 34:15 and continue for 80 min. A brief arousal to eyes-open, returning through 2A and 2B categories to a depth of 3B at 35:55, shows continuance of the sleep cycle.

Weighted total counts—Figure 9 (continuation

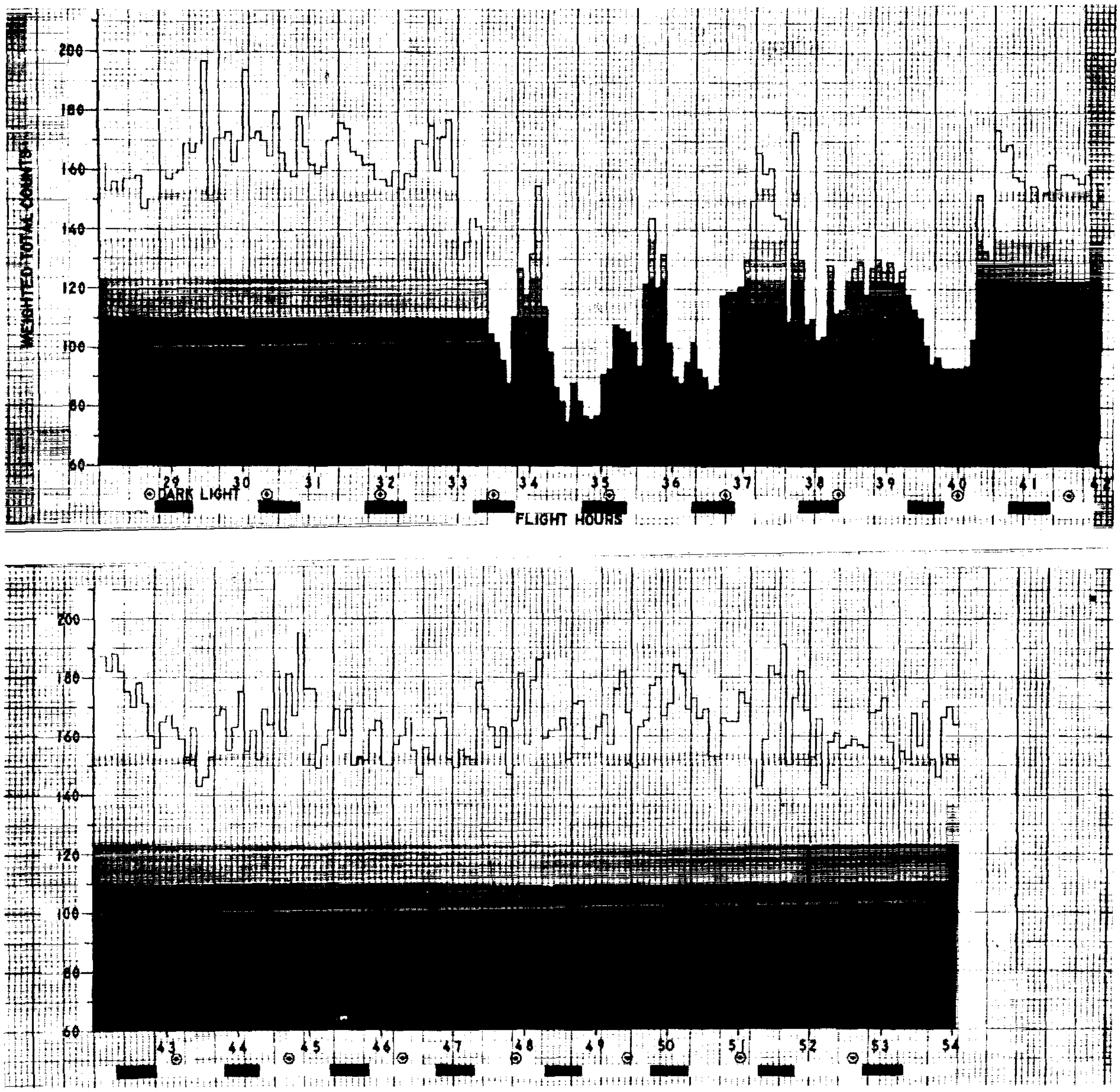


FIGURE 9.—Continuation of figure 4 between 28:00 and 54:00; the long sleep period between 33:00 and 41:00 is shown.

of fig. 4) shows the previously noted cycling by the state of consciousness continuing until the onset of the deep-sleep cycle at 33:00. We see now, quite elegantly displayed, six episodes of stage-IV sleep, with the first cycle at 33:45 lasting for 10 min and the third at 35:35 lasting only 5 min. The first and third cycles, with counts of 88 and 94, indicate a relatively light stage-IV sleep. The second cycle of sleep, just prior to 35:00, is the deepest stage-IV seen throughout the entire flight. All stage-IV cycles are relatively unstable during this time. The increased stability of the sleep profile in the last 25 percent of the sleep period argues for a return to a degree of neurophysiological homeostasis that had been absent until this time in the flight. Once again it is noted that the weighted total counts approach or enter the T_1-zone at times coincident with significant mission events.

Band parameters—Figure 10 shows slow-wave

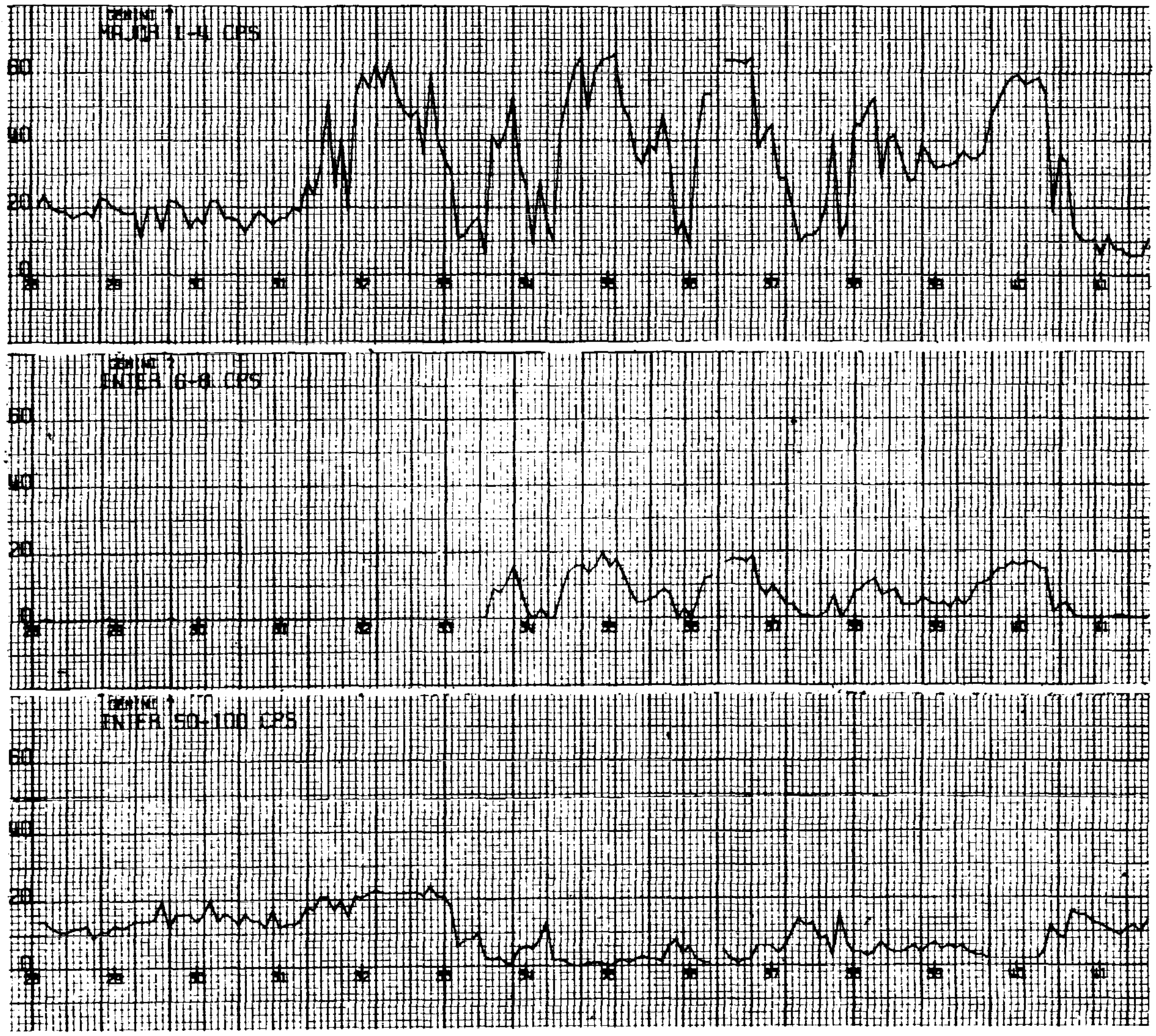


FIGURE 10.—Comparison of the major-period delta-index band-1 (1 to 4 Hz), the intermediate-period fast-theta-index band-3 (6 to 8 Hz), and the intermediate-period fast-activity band-10 (50 to 100 Hz).

(delta) activity from 31:30 to 37:15, but between 31:30 and 32:50 it is accompanied by fast frequencies shown in intermediate-period-band-10 (50–100 Hz) and reflects high activity. The slow-wave activity from 33:30 to 37:15 is accompanied by theta activity and reflects slow-wave sleep. The intermediate-period band-2 (4–6 Hz) differentiates these states very well, being high during sleep and absent during artifact.

From 36:00 to 42:00

Mission events—The flight plan schedules termination of the command pilot's sleep at 41:30. Medical reports (ref. 16) of the astronauts' subjective evaluation of their sleep during this mission correspond nicely with the quantitative evaluation of the analyzed EEG. The medical observations are so important that we quote them *in toto*:

> Neither crewman slept as soundly in orbit as he did on the earth, and this inflight observation was confirmed in the postflight debriefing. The pilot seemed to fall asleep more easily and could sleep more restfully than the command pilot. The

command pilot felt that it was unnatural to sleep in a seated position, and he continued to awaken spontaneously during his sleep period and would monitor the cabin displays. He did become increasingly fatigued over a period of several days, then would sleep soundly and start his cycle of light, intermittent sleep to the point of fatigue all over again.

Discriminant analysis—The fourth deep-sleep cycle of this period lasts for 20 min at the 4C-level, with lightening of sleep to the 3A-level for 15 min before waking to eyes-open at 37:15. After some 15 min of the eyes-open state there is progressive oscillation down to the light stage-IV sleep of category-4A. Reappearance of the 2B-category before 39:00 indicates that dreaming occurred during this span of moderate-to-light sleep. The sixth and final deep-sleep cycle of this period is seen as a staircase progression downward from 39:00 to 40:00, reaching deep stage-IV sleep of category-4C for 10 to 15 min at 40:00. All evidence of sleep is absent from this index after 42:00.

Weighted total counts—The depth of stage-IV sleep for the various cycles may be compared in the weighted total-count profile. The excellent stability of the stage-I plateau during the 38th and 39th hours is rather dramatic; the stability of the final stage-IV cycle around 40:00 reaches a plateau at 93 counts per second. Waking occurs at 40:20 with a slight rebound into stage-I sleep during the following 5 min, but with moderately high arousal that peaks after 42:00. Progressive decrease in the state of arousal culminates in total weighted counts of 143 at 43:30—close to stage-I sleep.

From 42:00 to 48:00

Mission events—A vision test was scheduled for 43:20 in the flight plan and another fuel test purge at 47:00. At some time during the 31st revolution (ref. 19), around 48:00, a launched Polaris missile was sighted by the astronauts. Anticipation of this event and the excitement of tracking an object under these dynamic circumstances must certainly have created a special state of vigilance and arousal.

Some time during the 3rd day of flight, a partial-phasing maneuver was directed (ref. 20), and a posigrade burn of 12.4 ft/sec was accomplished. This was a change in the mission's plan that took advantage of the excellent turnaround progress at the launch site in preparation for the next launching. This deviation from the flight plan and the subsequent maneuvering may have been related to recurrence of a state of vigilance.

The midline electrodes for the EEG became dislodged at 54:28. The left parieto-occipital electrodes had apparently become dislodged about 25 hours earlier (ref. 21).

Discriminant analysis—Eyes-open is the exclusive state except for one episode of T_1 from 43:50 to 46:05; the record then begins intermittent oscillation between categories T_1 and eyes-open that continues until the recording is terminated by dislodgment of the electrodes at 54:00. At 53:35 the dip to "2B" is artifact related to the last recorded meal.

Weighted total counts—Once again we see the state of consciousness entering into the T_1-zone in relation to the vision test and the fuel-cell purge. The oscillation of the weighted total count within and about the T_1-zone between about 45:40 and 47:20 may well have been related to the anticipation associated with the firing and tracking of the Polaris missile. The T_1-state seems to be closely related to events requiring special attention and high-level performance, as evidenced by the relation between known events during this flight and occurrence of this special electroencephalographic state.

Comparative Study

Having quantitatively analyzed the EEG recorded during this orbital flight and having interpreted these data in state-of-consciousness terms of sleep and alertness, we now turn to comparison with results from other studies. Such comparison helps fractionation of the portion of the observed disturbance of sleep that may be due to the following principal factors:

(1) The mission profile of 14-day, 24-hour confinement and a programmed work schedule in confined quarters

(2) The fact of deprivation of deep sleep during the command pilot's first "night" of sleep

(3) The degree of stress inherent in the extremely novel environment of orbital flight

It is important to determine what fraction of the disturbance seen in this flight profile is due to the orbital flight itself, in contrast with the fraction that may be due to anticipation and the circumstances of 14-day confinement in the capsule. In terms of strictly scientific "control," the ideal situation would be for the astronaut of this flight to simulate the flight without orbiting, so that the EEG might be recorded with all circumstances identical but for actual flight. Lacking this ideal if somewhat impractical control, we substitute a next-best set of data for comparative study.

A series of 14-day-night Apollo missions were simulated in the laboratories of Baylor University College of Medicine. The purpose was simulation as closely as possible of a three-man mission profile of the Apollo shots. One subject, a 25-year-old candidate for Naval Officers School, acted the role of systems engineer during one series of runs. A single-channel EEG, from essentially the same midline sites employed in the orbital flight, was recorded and analyzed on line by the same period-analytic techniques. While the entire 14-day run was analyzed, only the weighted total-count data of the first sleep period will be compared with the flight profile.

Figure 11 plots the weighted total counts in 5-min samples for the 9 hours of the first-night sleep which began at about 0100 hours. The subject had then been awake for some 15 hours, as had been the astronaut at the onset of his first sleep period in the 8th hour of flight. The extremely high value of 252 counts, 30 min before the onset of sleep, reflects a required exercise period. The "Apollo" sleep profile shows the following time-course attributes which may be compared to expected characteristics based on findings in large populations of experimental subjects (refs. 22 and 23).

The first stage-IV-sleep cycle occurs within the first 30 min from the onset of sleep; "no-stress" sleep is expected to show stage-IV slightly sooner. The first stage-IV-sleep cycle is the deepest cycle of the night; no-stress sleep also is expected to show deeper stage-IV in this first slow-wave cycle than in subsequent cycles. The three sequential slow-wave-sleep cycles tend to become progressively less deep throughout the night; no-stress sleep is expected to show such a trend, with three or four cycles expected. The deep-sleep cycles tend to stabilize at a given level for 20 to 30 min; no-stress sleep is expected to stabilize at a given level of sleep for somewhat longer periods. The last half of the sleep profile tends to develop a plateau at a relatively light level of stage-II sleep; no-stress sleep also is expected to attain a plateau in light sleep during the last third of the night.

A final transform of these data may be employed for evaluation of the sleep period. The lengths of time spent at various levels of sleep may be considered as percentages of total sleep time. Under this normalizing procedure we find the "Apollo" sleep profile to consist of the percentages indicated in table 4. The mean percentages of no-stress sleep (ref. 24) make it clear that, whatever stresses and subsequent disturbance of neurophysiological homeostasis may be associated with the "Apollo" mission, they are insufficient to produce more than moderate disturbance of the first-night sleep—manifested primarily by a marked reduction of stage-I sleep.

Let us now compare various episodes of sleep in the space-flight profile with the "base line" established from a simulated mission only.

From 08:20 to 10:15

Considering this period as the first 2 hours of a night's sleep, we find it highly unstable and extremely light, with absence of sleep of stages III and IV. This must be interpreted as a relatively severe disturbance of the sleep profile, as compared with both no-stress sleep and the "Apollo" data base. The percentages of the sleep levels in this episode show total deprivation of stage-III and -IV sleep, with extreme deprivation of stage-II.

From 14:20 to 15:40

This sleep episode shows an inversion of the expected time course, with each subsequent cycle becoming progressively deeper. The cycles do not stabilize for any considerable length of time, and only the third and last cycle reaches stage-IV sleep. This sleep episode shows increased percentages of all deeper stages at the expense of stage-I sleep and waking.

Interpreting this sleep episode as the first part of the night's sleep, we still find severe disturbance of the profile. One might contend that the

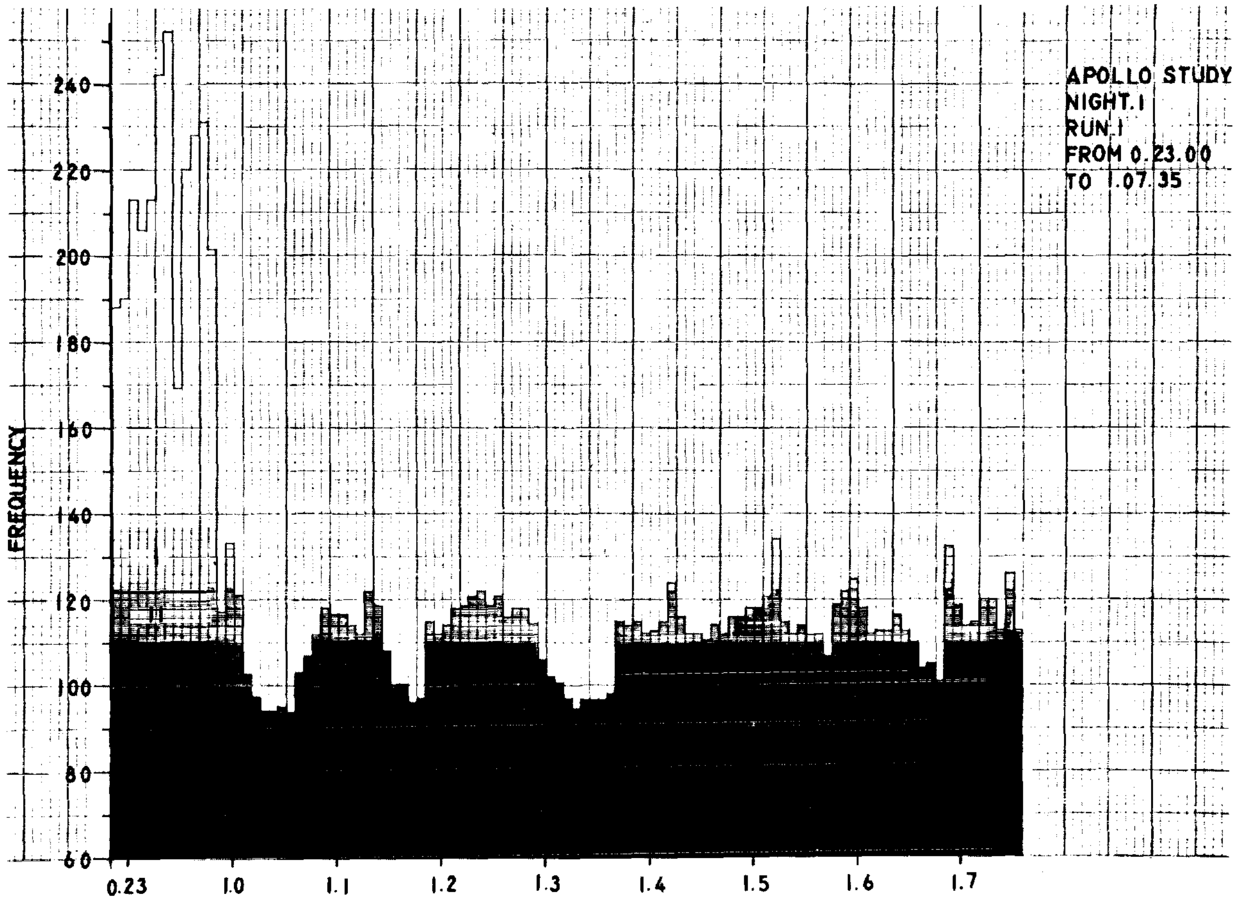


FIGURE 11.—Weighted total-count graph of the first night of the "Apollo" study.

time between 08:00 and 16:00 represents the total night's sleep that the astronaut would have experienced had he not been on this mission; under this construction the sleep profile is seen to be even more grossly disturbed than is indicated by the previous interpretations.

From 33:05 to 40:35

This sleep episode, the longest and deepest recorded on this flight, began about 40 hours after the astronaut's last ground sleep ended. Thus it would be perfectly reasonable to consider this to be his second "night's" sleep, but a more lenient interpretation of the degree of disturbance is allowed in continuance of the comparison with the first night of "Apollo" base-line data.

The first-stage-IV-sleep cycle occurs 45 min after the onset of sleep, 15 min later than in "Apollo"; this cycle is considerably lighter than the second or third cycle, and the relatively deep final cycle shows inversion of the expected sequence. The occurrence of six cycles of deep sleep is twice the degree of cycling of "Apollo." This degree of cycling is a disrupted profile probably produced by frequent arousals, movements, and changes in the stage-IV sleep (ref. 25). Only the final deep-sleep cycle shows the stability and duration that were present even in the sleep of "Apollo." The last quarter of the profile shows an initial tendency to develop a plateau at stage-II and stage-I sleep, but this tendency is disrupted by the unusual final deep-sleep cycle.

The percentages in this sleep episode show that the deep-sleep time has considerably increased from that in any previous portion of the flight, and that this increase is largely at the expense of

TABLE 4.—*Percentages of Time Spent Awake and in Four Stages of Sleep Under Four Conditions*

Awake or in stage of sleep	Normal no-stress sleep	"Apollo" night-1	Night after 2 nights of stage-IV deprivation	Orbital flight hours			
				08:20–10:05	14:20–15:40	08:20–15:40	33:05–40:35
Awake	0	0	0	57	30	71	14
Stage-I	29	7	23	39	25	15	19
Stage-II	49	63	50	4	25	7	19
Stage-III	8	15	7	0	10	4	17
Stage-IV	13	14	20	0	10	2	31

stage-I sleep which has been reduced to 19 percent. It may be argued that the degree of disturbance, represented by the above percentages, may be simply explained by the deprivation of deep sleep that has been so clearly seen during the first "night" of the flight. This explanation does not adequately explain the percentage levels of the command pilot's second "night." Experimental studies (ref. 26) show that not even 2 nights of deep-sleep deprivation produce the apparent "hunger" for stages III and IV seen in the flight. The percentages of the deep-sleep stages are grossly increased at the expense of stages I and II; the excessive waking time is of course further evidence of a disrupted pattern.

We must conclude that anticipation of the 14-day confinement and work schedule does not constitute enough stress to account for the disturbance in the first-night sleep profile, and that the deep-sleep deprivation of the first night does not explain the degree of disturbance of the sleep of the second night. It follows that the unique circumstances of an orbital mission must produce additional stresses which are more disruptive of the sleep profile than are these other factors. It is certainly known that the events and stresses of a day influence the subsequent sleep profile (ref. 27), and that simple real-life stresses such as the loss of a billfold or a husband's fight with his wife decrease the sleep percentage of stages III and IV during the subsequent night.

All-night sleep recordings from a given individual, over at least 3 or 4 nights (ref. 28), are required for quantitative determination of the degree of stress represented by certain percentage decreases of the various stages of sleep. We need multiple-night sleep recordings from the command pilot before we can evaluate the degree of the unknown stresses of orbital flight that account for the sleep profiles here reported. Sleep recorded during the early evening, rather than during the regular sleeping period, is not completely adequate for establishment of the "basal" sleep profile of a given individual. However, such base-line multichannel EEG's have been recorded from the command pilot both during sleep and under an active stimulus field; period analysis of these records should be performed to enable completion of the comparative studies and to extend the power of behavioral interpretations.

DISCUSSION

The command pilot's EEG covered the first 54 hours of orbital flight. From the point of view of biological information, this was an historic occasion, not only in marking the first recording of a continuous EEG from an orbiting American astronaut but also in providing the data base for the most extensive analysis in the history of electroencephalography. Almost every element in the total recording-playback system is special to this particular orbital flight: the sensors and recording devices, the reproduction and dubbing techniques, and the replay for analysis of copies of the master analog magnetic tape. The total system worked remarkably well, and even the components that failed (the electrodes that became dislodged) contributed the valuable negative information that more research and development was needed in this problem area of electrode configuration for long-term applications.

An event as unique as this flight is reported observationally rather than as a scientific experiment. This observational report emphasizes individuality but requires comparative studies for relation of the individual case to what is known from rigorously controlled experimental studies. For this reason, results of the Apollo simulation have been used for comparative interpretation. The depth of sleep and the temporal evolution of the sleep profile have been quantitatively measured by period analysis of the EEG both in this flight and in "Apollo," but further experimental work is especially essential for determination of the relation between stress in an individual and the consequent degree of disturbance in his sleep profile.

This study has strongly demonstrated the need for high-resolution documentation of current mission events during the flight. Period analysis of the EEG offers a high-resolution measure of alertness that can be translated into interpretation of stimulus-response appropriateness only if the stimulus field of the moment is known. The response of the alert individual—whether by task-performance or by EEG—should be appropriately related to the intensity or significance of the stimulus. For a high-intensity stimulus the EEG response should show high arousal; for a low-intensity stimulus it should show low arousal. For a stimulus requiring rapid high-level performance, the performance response should be commensurate. A stimulus allowing for relatively slow and approximate performance should evoke more-relaxed behavior. In brief, the response should be appropriate to the stimulus; it may be said that such a relation operationally defines good performance.

The EEG signals from the single- and double-channel placements have been shown to provide an index of arousal and of depth of sleep that has direct relevance to performance, but we are just beginning to exploit the full power of multichannel recording as a measure of adequate response. The multichannel EEG, with appropriate analysis for interrelating the activity between multiple sites, promises to be an extremely powerful tool for study of human behavior. Important for future long-term missions are the questions to be answered about continuing long-term interpersonal reactions between from three to five crewmen. The EEG promises to aid greatly understanding of the interactions and dynamics in a small human group; it may well be the only measure of response having adequate specificity to the stimulus as well as high enough resolution and sensitivity to allow for study and prediction of human interaction. Certainly no other single psychophysiological or psychological measure can be expected to suffice. After the fact it is quite clear that valuable scientific information was lost during this flight because the pilot was not instrumented for an EEG.

ACKNOWLEDGMENT

We thank both the disciplines and the individuals participating in this work, not in the usual sense of thanking but to make clear the degree of interdisciplinary effort that was required for its accomplishment. The results speak for themselves; in their scope and detail they are perhaps as unusual a contribution to the state of the art as was this orbital flight itself.

REFERENCES

1. AGNEW, H. W., JR.; WEBB, W. B.; AND WILLIAMS, R. L.: The First Night Effect: An EEG Study of Sleep. Psychophysiology, vol. 2, 1966, pp. 263–266.
2. BURCH, N. R.; AND GREINER, T. H.: A Bioelectric Scale of Human Alertness: Concurrent Recordings of the EEG and GSR. Psychiat. Res. Repts., vol. 12, 1960, pp. 183–193.
3. ADEY, W. R.; KADO, R. T.; AND WALTER, D. O.: Computer Analysis of EEG Data from Gemini Flight GT–7. Aerospace Med., April 1967, pp. 345–359.
4. MAULSBY, R. L.; FROST, J. D., JR.; AND GRAHAM, M. H.: A Simple Electronic Method for Graphing EEG Sleep Patterns. Electroencephalog. Clin. Neurophysiol., vol. 21, 1966, pp. 501–503.
5. PROCTOR, L. D.; ET AL.: The Analysis of Gemini–7 EEG Data by Three Different Computer Oriented Methods. NASA report under contract NSR–23–003–005.
6. MAULSBY, R. L.: Electroencephalogram during Orbital Flight. Aerospace Med., vol. 37, 1966, pp. 1022–1026.
7. KELLAWAY, P.; AND MAULSBY, R. L.: M–8 Experiment: Inflight Sleep Analysis Gemini–VII Flight. NASA preliminary report, Feb. 7, 1966.
8. BURCH, N. R.; AND CHILDERS, H. E.: Information Processing in the Time Domain. Information Storage and Neural Control. Charles C. Thomas, Springfield, Ill., 1963.
9. SALTZBERG, B.; AND BURCH, N. R.: A New Approach to Signal Analysis in Electroencephalography. I.R.E. Trans. Med. Electron., 1957, p. 24.
10. COOLEY, W. W.; AND LHONES, P. R.: Multivariate Procedures for the Behavioral Sciences. John Wiley and Sons, Inc., 1962.

11. JOUVET, M.: Paradoxical Sleep—A Study of Its Nature and Mechanisms. Sleep Mechanisms. Progr. Brain Res., vol. 18, 1965, pp. 20–62.
12. BERRY, C. A.; COONS, D. O.; CATTERSON, A. D.; AND KELLY, G. F.: Man's Response to Long-Duration Flight in the Gemini Spacecraft. NASA SP–121, vol. 25, 1966, p. 235.
13. GRISSOM, V. I.; MCDIVITT, J. A.; COOPER, L. G.; SCHIRRA, W. M.; AND BORMAN, F.: Astronauts' Reactions to Flight. NASA SP–121, vol. 27, 1966, p. 272.
14. MIGLICCO, P.; COHEN, R.; AND DEMING, J.: Electrical Power and Sequential Systems. NASA SP–121, vol. 6, 1966, p. 54.
15. HODGE, J. D.; AND ROACH, J. W.: Flight Control Operations. NASA SP–121, vol. 20, 1966, p. 184.
16. BERRY, C. A.; COONS, D. O.; CATTERSON, A. D.; AND KELLY, G. F.: Man's Response to Long-Duration Flight in the Gemini Spacecraft. NASA SP–121, vol. 25, 1966, p. 241.
17. FOULKES, D.; AND VOGEL, G.: Mental Activity at Sleep Onset. J. Abnormal Psychol., vol. 70, no. 4, 1965, pp. 231–243.
18. ANON.: Mission Program Profile of the Gemini–VII Flight. NASA–S–66–170 (fig. 7.1.1).
19. BRENTNALL, B.: Experiment D4/D7, Celestial Radiometry and Space-Object Radiometry. NASA SP–121, vol. 38, 1966, p. 374.
20. HODGE, J. D.; AND ROACH, J. W.: Flight Control Operations. NASA SP–121, vol. 20, 1966, p. 186.
21. KELLAWAY, P.: Experiment M–8, Inflight Sleep Analysis. NASA SP–121, vol. 45, 1966, p. 424.
22. DEMENT, W.; AND KLEITMAN, N.: Cyclic Variations to Eye Movements, Body Movements, Body Motility and Dreaming. Electroencephalog. Clin. Neurophysiol., vol. 9, 1957, p. 673.
23. WILLIAMS, H. L.; HAMMACK, J. T.; DALY, R. L.; DEMENT, W. C.; AND LUBIN, A.: Responses to Auditory Stimulation, Sleep Loss and the EEG Stages of Sleep. Electroencephalog. Clin. Neurophysiol., vol. 16, 1964, pp. 269–279.
24. WILLIAMS, R. L.; AGNEW, H. W.; AND WEBB, W. B.: Sleep Patterns in Young Adults: An EEG Study. Electroencephalog. Clin. Neurophysiol., vol. 17, 1964, pp. 376–381.
25. MONROE, L. J.: Psychological and Physiological Differences between Good and Poor Sleepers. J. Abnormal Psychol., vol. 72, no. 3, 1967, pp. 255–264.
26. AGNEW, H. W.; WEBB, W. B.; AND WILLIAMS, R. L.: The Effects of Stage Four Sleep Deprivation. Electroencephalog. Clin. Neurophysiol., vol. 17, 1964, pp. 68–70.
27. LESTER, B. K.; BURCH, N. R.; AND DOSSETT, R. G.: Nocturnal EEG-GSR Profiles: The Influence of Pre-Sleep States. Psychophysiology, vol. 3, no. 3, 1967, pp. 238–248.
28. WILLIAMS, H.; AND WILLIAMS, C.: Nocturnal EEG Profiles and Performance. Psychophysiology, vol. 3, no. 2, 1966, pp. 164–175.

CHAPTER 9

ANALYSIS OF ELECTROENCEPHALOGRAPHIC DATA FROM ORBIT BY THREE DIFFERENT COMPUTER-ORIENTED METHODS

L. D. Proctor

The problem in assessment of electroencephalograms (EEG's) has been the subjective element in visual evaluation of such records, as well as the need for extensive training (2 to 3 years) before a suitable candidate can perform this task. There have been a variety of computer-type analyses of EEG's during the last decade, and we are still searching for a reliable, inexpensive, on-line, data-processing procedure. We are aware that our Soviet colleagues routinely assess EEG's by computer, using relatively inexpensive hardware.

Here I endeavor to show how four different assessment procedures (visual assessment, zero-crossings technique, smoothing and peak-counting technique, and Weibull statistic) may be utilized. The advantages and disadvantages of each will also be indicated.

I wish to point out that, in the space-flight analog records of EEG's supplied to Henry Ford Hospital, channel-II (midline-central to midline-occipital electrodes) could not be assessed reliably after about 26 hours of flight (26:00). Channel-I (left-central to left-occipital electrodes) provided the more assessable EEG until about 54:00. These channels during the first 26 hours of flight showed no significant difference by our four assessment procedures. Therefore, this report will be limited to the findings from analyses of channel-I during its approximately 54 hours of recording.

ZERO-CROSSINGS TECHNIQUE

The zero-crossing method for reduction of electroencephalographic data is a simple pattern-recognition program that arbitrarily defines the frequency components of the EEG but disregards the amplitude of the signal.

The EEG is recorded in analog form on magnetic tape before the tape is played back through a digitizing system. Two such systems were used: an IBM system produced a digitized record of 312.5 points per second of real-time recording; an ASI (Advanced Scientific Instruments) system yielded 242 points per second of real time. In both cases the rule-of-thumb requirement of 5 points per highest frequency analyzed was met since we had agreed to ignore all frequencies above 50 Hz in the EEG. Both digitizing systems had sufficient core memory to permit an entire EEG record to be digitized without interruption and subsequent loss of any digits; in both, the range of the scale value of the digits was from 0 to 510. The initial base-line value for the EEG is chosen at the midpoint of the range (in this case at a value of 255). This base line is subsequently altered as follows: The values between a present upward crossing of the base line and the previous upward crossing are averaged. The value of this recent upward crossing is then subtracted from the average. This difference is then multiplied by 0.05, and the result is added to the present base-line value. With use of this method the arbitrary base line tends to follow the drifts in the average value of the signal. The small weighting factor of 0.05 was used so that the base line would follow only substantial and consistent shifts.

As the values of the digitized EEG cross the base line, the time between each pair of sequential crossings is measured and stored in two sets of registers. In one set it is stored as a half-wave of a particular frequency (0.5 to 50 Hz) depending on the time between crossings; in the other set the amount of time is stored. Thus there are two sets of registers: one for frequency counts and one for total time per frequency for a given sample of EEG. Each set contains 51 individual registers (one register for each frequency from 0.5 to 50 Hz). From these registers we extract the frequency counts and the total time present in each frequency band that has been established in clinical use; that is, delta (0.5–3 Hz), theta (4–7 Hz), alpha (8–15 Hz), beta-1 (16–25 Hz), beta-2 (26–35 Hz), and beta-3 (36–50 Hz).

PARAMETRIC ANALYSIS OF THE SPACE-FLIGHT EEG'S

Even cursory visual examination of any electroencephalographic record reveals, if nothing else, the extreme variability of the recorded potentials within a given subject. When the data are summarized by automatic methods such as zero-crossings (Z/C) or smoothing and peak counting (SPC), which hopefully eliminate intuitive interpretations, the problem becomes particularly acute and one must define a standard or "normal" population against which the reliability of a given signal change may be estimated. That is, are the correlations of fluctuations in an electroencephalographic signal truly representative of a change in the physiological status of the organism, or simply random variations independent of the behavior in question? Ideally one would desire: (1) a library of data, taken from many subjects under standardized conditions, that would provide maximum generality and a broad base for comparison; or, alternatively, (2) repeated measures taken from the same subject that, although restricted to one individual, would provide somewhat greater power for comparative techniques. However, the ideal is not often realized, and the present analysis was forcibly restricted to less-desirable alternatives.

Thus, while the entire raw EEG record was reduced by the methods described, the parametric analysis dichotomized the record into two separate epochs which respectively provide (1) a standard population (first 7 hours of flight) and (2) estimates of the subject's state of consciousness during the remainder of the flight. The first 7 hours were chosen as the standard not only because the EEG signal was extremely stable during this period, but also because it seemed reasonable to assume that the subject was in an alert, active state during this time rather than during the relatively routine hours later in the flight. It must be emphasized, though, that this procedure incorporated several drawbacks, not the least of which was a biased standard that would not be present if preflight EEG's were available.

Because of the correlation between individual means and their respective variances, the parametric analysis transformed each raw score [frequency count over successive 15-sec intervals (Z/C) or 1 min intervals (SPC)] according to the following relation (ref. 1):

$$X' = X + 0.5$$

where X is the raw score from either Z/C or SPC, and X' is the transformed score used in subsequent analysis.

This transformation was carried out for homogeneity of variance over the entire recorded period, and all remaining parametric calculations were performed on these transformed scores. The standard epoch (first 7 hours of flight) was then ordered into four separate populations according to classical delta, theta, alpha, and beta frequency-band criteria, with each population described by the mean and standard deviation. The remaining flight record was then analyzed in blocks of 15-min (Z/C) or 10-min (SPC) periods with the mean amount of activity occurring in each of the four frequency bands of these periods being compared to the mean and standard deviation calculated for the standard epoch.

Z/C METHOD OF DATA REDUCTION

Figure 1 presents the mean activity computed for each of the analyzed frequency bands over 41 hours of the flight. Due to noise and extreme fluctuations of the signal, the subsequent record was not analyzable by the Z/C method.

Representing the initial step in the parametric analysis, figure 1 provides an overall visual presentation of the electroencephalographic activity and demonstrates in a general manner how the indicated overt activities of the subject produce changes in the recorded potentials. For example, the stable first 7 hours of flight (our standard epoch) contrast with the 8th hour when

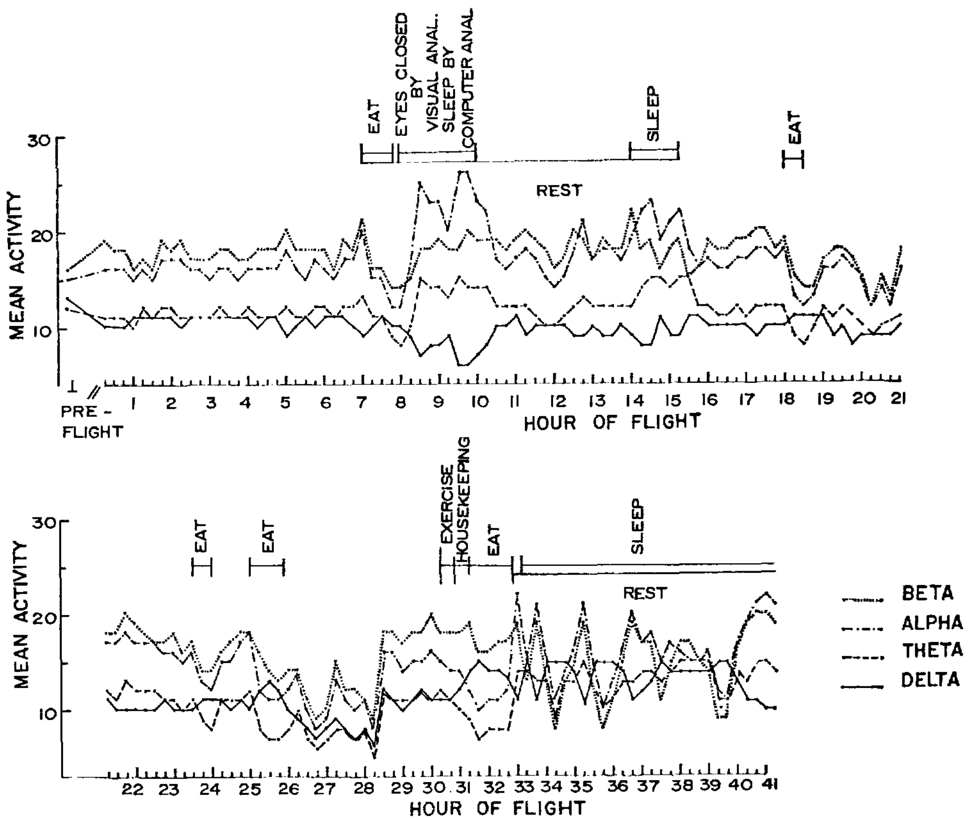


FIGURE 1.—Mean activity computed for each of the analyzed frequency bands over the first 41 hours of flight; raw output from zero crossings.

the astronaut was relaxed with eyes closed and effecting (as expected) substantial increase in recorded alpha activity. Due to the manner in which the individual frequencies are derived from base-line crosses (see description of Z/C methods), the quality of the EEG signal may be inferred since any sharp decrease in quantity of activity of all bands above delta is indicative of off-scale potential shifts or large amounts of muscle activity (note particularly the indicated eating periods) carrying the signal away from the base line. Thus, while figure 1 is quite general, it does indicate that gross behavioral states are indeed reflected in the automatic analysis of the recorded signal.

With the data summarized in this manner one can specifically compare the changes in amounts of activity that occurred after lift-off by simple *t*-tests for differences between the mean activities recorded during the 10-min period prior to lift-off and during the first 10 min of flight. This comparison is summarized by a histogram (fig. 2) showing as individual bars the mean activity within each of our defined bands and also the results of the applied statistic. There is a significant decrease in delta activity recorded during the initial 10 min of flight from that recorded before lift-off, while alpha and beta exhibit significant increases for the same comparison; change in theta activity was insignificant. It must be noted, however, that a large portion of the lift-off period is characterized by extreme-muscle artifact that could conceivably account for a goodly portion of the increase in the high-frequency (beta) activity.

For determination of the subject's state of consciousness over the recorded period of the flight, the parametric analysis proceeds by initial calculation of the mean activity in each of the frequency bands during successive periods lasting approximately 12 min. The variability in the lengths of the individual periods and the number

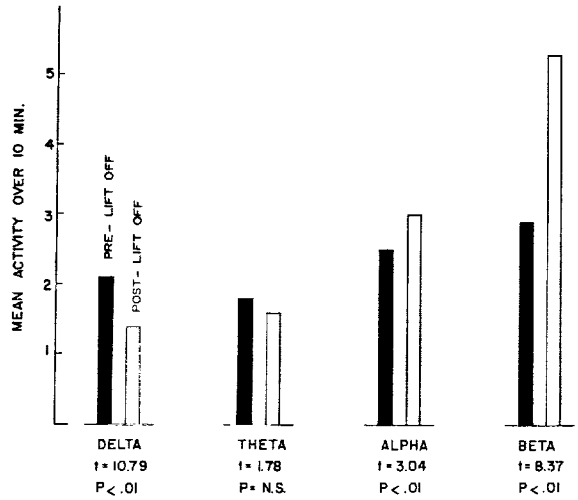


FIGURE 2.—Comparison of mean activities during 10-min periods preceding and following lift-off, and results of the applied statistic.

of periods in each hour result from change in quality of the signal and consequent rejection of data. The individual means are then presented in figure 3 as they occur outside the bounds of ± 1 standard deviation of the standard epoch. Thus figure 3 shows the occurrence of activity that deviates from what we have defined as activity recorded when the subject is in a highly alert and active state.

With the EEG activity depicted as in figure 3, one can proceed to more detailed analysis and determination of the various stages of sleep (ref. 2) by evoking the following definitions:

(1) Stage-0—if all activity is within 1 standard deviation of standard epoch; and if only alpha activity exceeds 1 standard deviation

(2) Stage-I—if alpha and theta mean activities exceed 1 standard deviation of standard epoch

(3) Stage-II—if only theta mean activity exceeds 1 standard deviation of standard epoch

(4) Stage-III—if both theta and delta activities exceed 1 standard deviation of standard epoch

(5) Stage-IV—if only delta mean activity exceeds 1 standard deviation of standard epoch

Figure 3 represents the results of application of these criteria to the data as presented in figure 2. Each figure indicates the subject's state of consciousness during 12-min periods by the vertical extension of the individual bars, and completely agrees with visual analysis of the electroencephalographic sleep records. It should be noted that the sleep period depicted in figure 4 corresponds to the results of the visual analysis (ref. 3) and is in disagreement by about 4 hours with the initial actual sleep period indicated by the flight log supplied to Henry Ford Hospital along with the analog tapes of the flight. However, the visual analysis does indicate that there is an "eyes-closed" period at the initiation of the rest period, but that actual sleep did not occur until about 14:00, with the intervening period characterized by relative wakefulness. Our parametric analysis, on the other hand, indicates that the subject did sleep very lightly (stage-I only) as shown by the flight log (see fig. 1), but the period was classified as "eyes closed but awake" by visual criteria.

The parametric analysis has been applied with varying success to both of the described methods of data-reduction: Z/C and SPC. While the Z/C technique provided the most useful and sensitive data for parametric analysis, one should also note that the standard-deviation criteria were originally set out for this method without regard to the characteristics of the SPC analysis. At least for the present there seems to be no *a priori* reason to assume that a similar parametric analysis could not be conceived for SPC once the program for various parameters were established from a sample of data considerably larger than from the recording of one session from one individual. However, parametric analysis and its benefits can only be applied usefully to Z/C data with realization of the shortcomings of the Z/C technique, particularly with regard to analog records of poor quality.

Despite the problems inherent in the simple Z/C analysis (inability to use poor-quality records such as those derived during eating), the method in combination with our standard-deviation criterion has provided extremely high resolution of the subject's state of consciousness as recorded during the flight before the signal began to deteriorate at about 42:00. The benefits of this analysis are further demonstrated by its ability to classify extremely light sleep (for example, that occurring between 08:00 and 10:00) which is classified as "eyes closed but awake" by routine visual analysis. These facts permit speci-

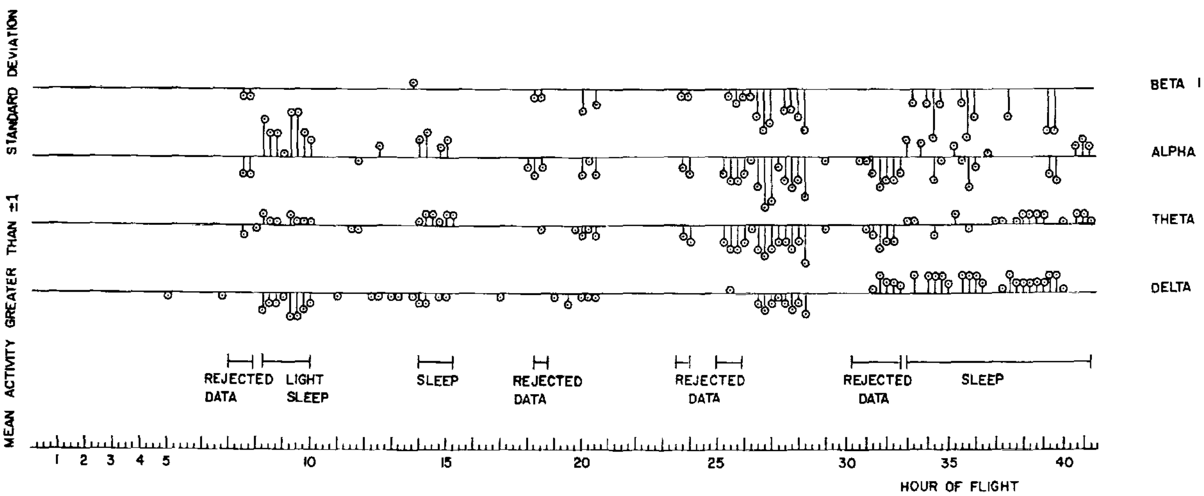


FIGURE 3.—Standard-deviation analysis with Z/C data.

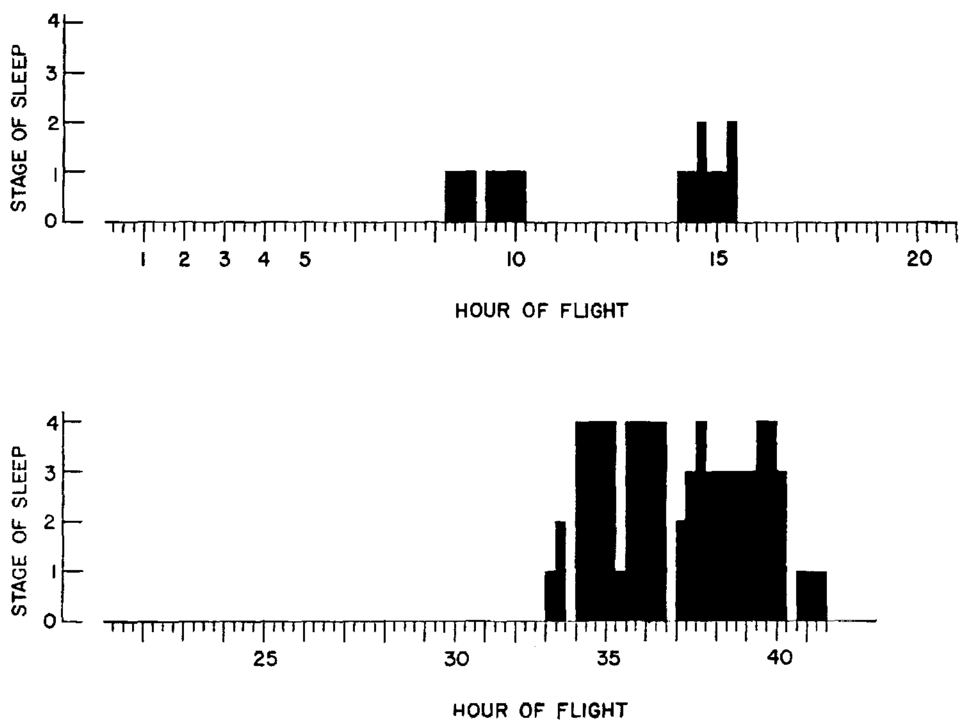


FIGURE 4.—Degree of consciousness from zero crossings and standard-deviation criterion.

fications of reliability of the EEG signal in terms of probability of occurrence of "EEG"-indicated behavioral states, and have provided the electroencephalographer with a quantitative tool heretofore unrealized in the clinical evaluation of EEG's.

Finally it should be emphasized that evaluation of the EEG signal by Z/C is independent of the individual reading the record since the analysis is divorced from any subjective interpretations of the signal. This system (Z/C and standard-deviation criterion) provides mutually exclusive categories incorporating the full range of EEG frequencies, the resultant benefit being a substantial reduction in ambiguity as to the state of consciousness of the subject. Approximate real-time analysis of electroencephalographic indicants of states of consciousness for any individual capable of reading numerical computer printouts is now available.

DIGITAL SMOOTHING AND PEAK-COUNTING TECHNIQUE

Let us consider one channel of "digitized" electroencephalographic signal: a series of positive integers proportional to the value of the signal voltage sampled at regular intervals. The process basic to the analysis is identification of maxima (peaks) and minima (valleys) in the digitized signal. A maximum is identified by a series of three points, the middle of which is greater than the other two; for a minimum, the middle point must be less than the other two. (When several successive points have the same value, and these points are followed and preceded by points both of which are less or greater in value, then a maximum or minimum, respectively, is said to have occurred at the middle value.) A pair of successive maxima define the boundaries of a valley wave, and a pair of successive minima define a peak wave (fig. 5). We are concerned with four properties of individual waves: frequency, amplitude, symmetry, and complexity. In addition we may consider small groups or sequences of waves according to conditional probabilities or in terms of larger patterns.

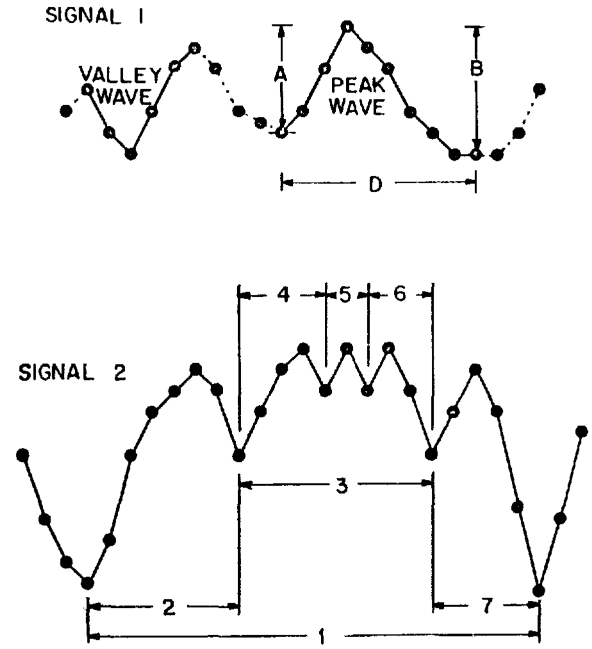


FIGURE 5.—Illustrations of wave properties.

Figure 5 illustrates the definitions of the various wave properties. Signal-1 includes valley and peak waves. Note that the right-hand minimum, defining the peak wave, is the middle point of three equal values. For the peak wave, the frequency is the reciprocal of the duration (D) in seconds; its amplitude is the average of A and B, and its symmetry is the ratio A:B (or B:A, whichever is larger). Similar definitions apply to the valley wave.

It is well known that, in the EEG signal, waves of different frequencies may be present simultaneously; fast waves are superimposed on slow waves. Signal-2 (fig. 5) illustrates this characteristic of complexity—waves 4, 5, and 6 are superimposed on wave-3; and waves 2, 3, and 7 are superimposed on wave-1. If appropriate low-pass electronic filters were applied prior to digitizing of the signal, subsequent analysis might yield only waves 2, 3, and 7; if a low enough filter were used, analysis might yield only wave-1. We may define simple waves as those that are identifiable without filtering and complex waves (waves 1 and 3 in fig. 5) as those requiring filtering for identification. Complete specification of the complexity of a particular wave would require knowledge of all waves (revealed by all possible filters) that overlap in time with the given wave.

In the specific analyses described below, a process of digital smoothing by computer is performed, instead of electronic filtering, to reveal this property of complexity. In the first of several stages (arbitrarily chosen) of smoothing, all waves are identified and those greater than 50 Hz in frequency are replaced by points of constant value so that they will not affect the identification of maxima and minima. (The smoothing process is described in detail below.) The new maxima and minima are determined, and the wave properties are computed. The frequencies, amplitudes, and symmetries determined after 50-Hz smoothing constitute one set of data. Similarly the smoothed signal is now resmoothed so that any waves having frequencies greater than 25 Hz are replaced by constant values. The wave properties are again determined, and a set of data for 25-Hz smoothing is obtained. Similarly sets of data for 15- and 8-Hz smoothing are determined.

The details of the smoothing process are important, and many other methods were tested before the present one was chosen. The most obvious method might be some type of averaging process; such methods are unsatisfactory because they provide a very unpredictable frequency cutoff and may lead to generation of frequencies not present in the signal because of phase relations (known as "aliasing" in power-spectrum theory). Another problem is that a high-amplitude, short-duration spike may be counted, after averaging, as a low-amplitude slow wave. Another possibility is simply use of successive stages of electronic filtering; for many applications this may be adequate, but it is probably less convenient, and some of the problems associated with averaging also apply to electronic filtering.

In the smoothing process described below, a peak or valley wave is replaced by points of constant value if it is of greater frequency than a specified cutoff. Before the present method was chosen, several variations were rejected on logical grounds (often after testing of revealed weaknesses) such as smoothing on the basis of criteria for half-waves, use of a replacement process in which the constant value runs from minima to minima or from maxima to maxima (see below), and smoothing of only peak waves.

The procedure used for digital smoothing is as follows: If a peak or valley wave is of greater frequency value than the specified cutoff, it is smoothed; but if it is of smaller frequency, it remains in its original form. As long as the previous wave has not been smoothed, we consider every (overlapping) peak and valley wave. If a peak wave is smoothed, the next (adjacent) peak wave (the overlapping valley wave being disregarded) is considered. Correspondingly if a valley wave is smoothed, the adjacent valley wave is next considered.

By smoothing we mean replacement of points

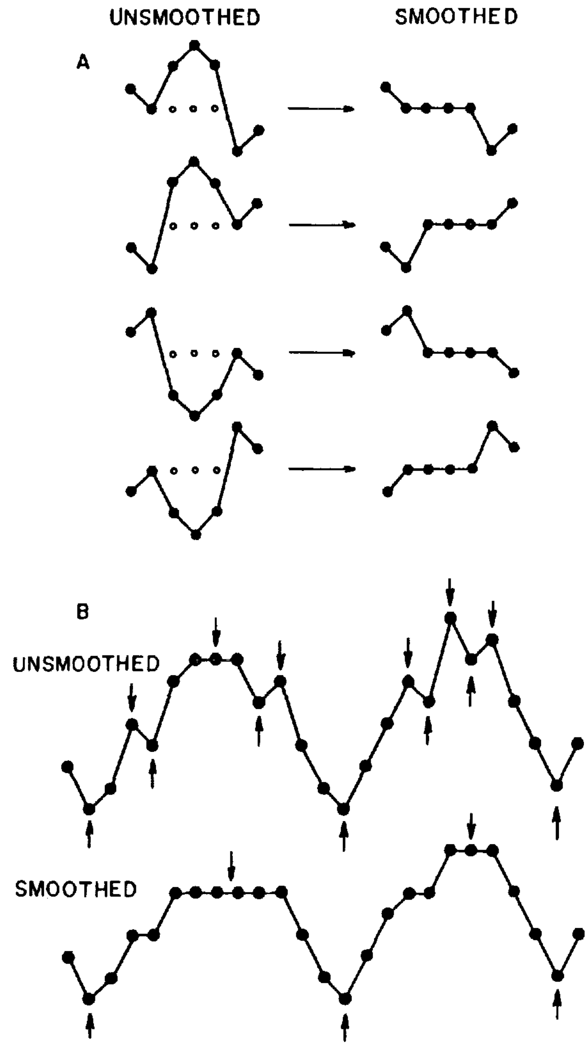


FIGURE 6.—Illustrations of the digital smoothing process.

according to one of the four procedures illustrated in figure 6A. If a peak wave is smoothed, we replace the points shown by values equal to the greater of the two minima bounding the wave. Note that the replacement begins at the greater minima and proceeds across the wave until the wave boundary is reached. Valley waves are smoothed by replacement according to the smaller of the two maxima. It is important to note that, if the initial smoothing criterion is chosen at too low a frequency, all waves above the criterion frequency may not be smoothed. Smoothing must either be repeated several times at a low frequency or be done by use of a series of successively lower stages. To clarify this point let us say that we are initially smoothing at 20 Hz. Two 70-Hz waves may be smoothed to form a 35-Hz wave which is above the cutoff. If the smoothing process is repeated, the 35-Hz wave disappears.

Figure 6B shows the smoothing process applied to a hypothetical signal. Six maxima and seven minima are shown (by arrows) for the unsmoothed signal, but for the smoothed signal there are two maxima and three minima.

A feature of the pattern-analysis is the ease with which certain types of nonelectroencephalographic signals can be recognized and rejected. Muscle potentials appear in the form of high-amplitude, high-frequency waves. Appropriate amplitude and frequency criteria can be established, and, if a given number of such waves are found within a period of time, that period can be disregarded. Another type of signal easily recognized as nonelectroencephalographic is a number of consecutive points of extremely high or low value; such signals are characteristic of recordings taken with loose electrodes. Illustrations of artifact-rejection are presented below.

ANALYSES BY DIGITAL SMOOTHING AND PEAK-COUNTING TECHNIQUE

Several different versions of the analysis described in the General Method section were applied to the flight-EEG data (channel-1). The first version to be discussed is a complete categorization of the waves according to frequency, amplitude, symmetry, and smoothing cutoff; it is applied to several representative epochs of the signal. This analysis is presented first so that the reader will understand better the general method and the more-summary analyses that follow. These summary analyses will cover the total amount of valid EEG signal.

Analysis-1—In this analysis, EEG waves were categorized according to 10 ranges of frequency, three ranges of amplitude and symmetry, and four criteria for smoothing. Frequencies greater than 50, 25, 15, and 8 Hz were successively smoothed out of the signal. For each smoothing cutoff, valley and peak waves were classified as follows: The frequency bands were 0 to 0.99 Hz, 1.00 to 2.99 Hz, 3.00 to 4.99 Hz, 5.00 to 6.99 Hz, . . . 15.00 to 16.99 Hz, and greater than 17.00 Hz. The amplitude categories were 1 to 19, 20 to 79, and 80 units or greater. The units come from an arbitrary scale of 0 to 510, which was used for digitization of the signal. One hundred units, the amplitude of a large alpha wave, was about 50 μV. Symmetry was defined above with reference to figure 5; the measure of symmetry used was whichever of the ratios $A:B$ and $B:A$ that exceeded 1. The categories for symmetry were 1.000 to 1.200 (symmetric), 1.201 to 3.000 (asymmetric), and greater than 3.000 (very asymmetric).

Five 10-sec epochs of EEG (fig. 7), each representing a stage of sleep [as interpreted by Maulsby (ref. 3) and verified by me] from the second flight-sleep period, were compared. They were chosen for being relatively homogeneous throughout the 10-sec period because they contained no artifacts by the criteria described and because they fitted clearly into their respective categories.

Since there are so many categories to be considered and because it is desirable to see all the categories together, it was decided to represent the number of waves falling into each category by a circle of which the diameter is proportional to the number. In figure 8, grouping is by sets of nine categories (3×3 arrays) in which amplitude increases from left to right and asymmetry increases from top to bottom. Thus the upper-right corner of a 3×3 array represents an almost asymmetric wave of high amplitude. The lower-left corner would be a highly asymmetric wave of low amplitude. It is apparent from figure 8 that most of the activity falls in the middle-amplitude, medium-asymmetry category (all circles except the middle are filled for emphasis of deviations from the middle ranges).

In examining figure 8 it is well to keep several ideas in mind. The first point is that certain bands must have values of zero since smoothing is on the basis of frequency; for example, 8-Hz smoothing removes all alpha activity from the signal. Another point is that at high smoothing cutoffs only simple waves (those without higher frequencies superimposed on them) are identified. It was found, for example, that all delta activity (1–3 Hz) is complex since the 50- and 25-Hz smoothing categories are all zero. (The 50-Hz smoothing category is not shown since it is very

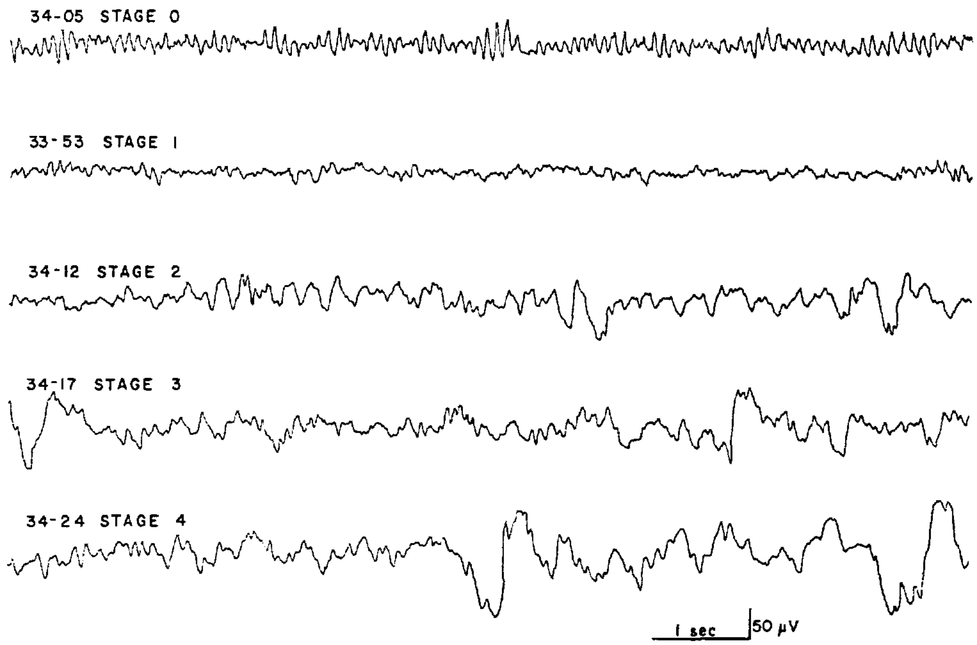


FIGURE 7.—Five 10-sec epochs of EEG, each representing a stage of sleep used for the analysis illustrated by figure 8.

similar to the 25-Hz category.) The third point is that for slow waves the lower-amplitude categories have little meaning in terms of the ordinary visual analysis of EEG's. When a signal is progressively smoothed at lower and lower frequencies, slow waves appear even if they represent only some kind of statistical fluctuation or possibly "spindling." If we consider only higher-amplitude categories for the slower waves (at low-frequency, smoothing cutoffs), however, these statistical fluctuations disappear. (One should note that, although we do not readily "see" these fluctuations, they are still valid properties of the signal and may turn out to be of interest.) The last point is that in examination of this figure a given number of waves in the low-frequency categories represents a greater percentage (in terms of time) of the signal than does an equal number of high-frequency waves.

Let us first consider the obvious differences between the EEG epochs characterizing the various stages of sleep, disregarding symmetry for the moment. State-0 is on the borderline between waking and sleeping (resting with eyes closed) and is characterized by a strong alpha rhythm (around 8 to 13 Hz). Figure 8 shows large amounts of 9- to 11-Hz and 11- to 13-Hz activity at all smoothing cutoffs. Stage-I of sleep is characterized by its low-voltage activity with complete lack of spindling. The categories for stage-I show that the activity is spread through the theta, alpha, and beta bands, with complete absence of high-amplitude waves (no circles in the right-hand columns of the 3×3 arrays); the highest numbers of frequencies greater than 17 Hz appear in this stage. Most characteristic of stage-II sleep is moderately high-voltage theta activity; we see the reappearance of high-amplitude activity in figure 8 under the 3- to 5-Hz (low theta) category. Stages-III and -IV of sleep are characterized by increasing amounts of high-voltage delta waves. The chart shows increasing amounts of high-amplitude activity in the 1- to 3-Hz range; in addition there are relatively large values in the 13- to 17-Hz ranges that may be related to the sigma (14-Hz) activity usually associated with stage-III.

This brief examination of figure 8 may have shown that for some bands there is a relatively constant number of waves for all stages. Most

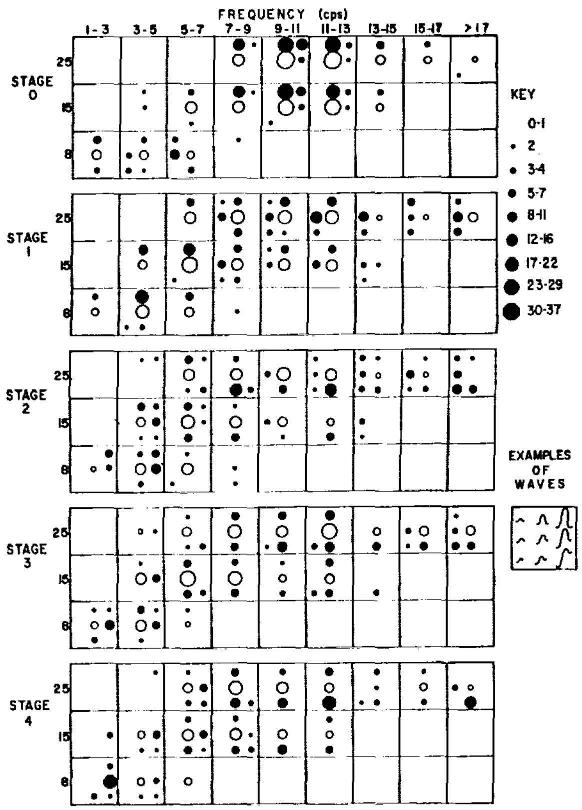


FIGURE 8.—Analysis of the five 10-sec epochs shown in figure 7.

constant perhaps is the 5- to 7-Hz band for 15-Hz smoothing. This band only shows a significant decrease for stage-0. Other bands in the middle-frequency ranges and at middle smoothing cutoffs tend to remain fairly constant. This constancy probably reflects the "random" character of EEG signals. In most cases there will always be a certain number of waves in the middle categories.

The top rows of the 3×3 arrays indicate highly symmetric waves; the bottom rows, highly asymmetric waves. As might be expected, the stage-0 alpha rhythm shows the greatest number of symmetric waves, and high-amplitude alpha is almost entirely symmetric. In stages-I and -II, low-frequency theta waves (3 to 5 Hz) seem to be relatively symmetric. More striking perhaps is the large number of asymmetric waves that appear occasionally, such as in the alpha range for stages-II and -IV. Asymmetry was particularly high for the 11- to 13-Hz band at the 50-Hz smoothing cutoff (not shown in fig. 8). Apparently there is a large amount of this asymmetric alpha activity superimposed on the slower delta and theta waves.

The complexity of the EEG signal may be determined by examination of changes in the distributions of waves after successive smoothings. Let us first consider stage-IV because it is apparent even from casual observation of an EEG chart that the slow delta waves have many higher frequencies superimposed on them. At 50-Hz smoothing (not shown in fig. 8) there are almost no waves below 7 Hz, but large numbers appear between 7 and 13 Hz and above 17 Hz. At 25-Hz smoothing, almost all the activity above 17 Hz has disappeared, and waves around 8 Hz are appearing. This shift to lower frequencies continues until, at 8 Hz, high-amplitude 1- to 3-Hz waves predominate. Within the 9- to 11-Hz column the number of waves decreases from the 25- to 15-Hz cutoff even though the waves are slower than the criterion frequency; an example will make clear how this can occur.

Consider two adjacent "peak waves," one 10-Hz and one 20-Hz, with the middle minima much higher than the two outer minima. With a 15-Hz criterion, the 20-Hz wave is smoothed, but the 10-Hz wave is not. However, when the maxima and minima are redetermined, the 10-Hz wave also disappears, and we are left with one 6.7-Hz wave, the combination of the 10- and 20-Hz waves.

Compare the smoothing process for stage-IV with that for stage-0. With a fairly homogeneous pattern such as the alpha rhythm, the 9- to 13-Hz activity is mostly retained even down to 15-Hz smoothing. However, for stage-0 at 8-Hz smoothing, some low-amplitude delta waves are recorded; these are due to slow shifts in the overall alpha pattern and to spindling. These "alpha associated" delta waves are easily distinguished from stage-IV sleep deltas by their low amplitude.

Analysis-2—In analysis-2 there is no categorization according to amplitude and symmetry. With the limitation of no amplitude categories, the results are more difficult to relate to those obtained by visual evaluation, and certain types of changes in the signal are not apparent. In analysis-3 it is demonstrated that, with the application of simple amplitude considerations, these limitations are no longer in effect.

Details follow of the computer program used: Epochs of 10.0 sec were regularly sampled every minute of the entire recording of valid EEG. For each epoch, all valley and peak waves were determined for each of the four smoothing cutoffs used for analysis-1. For each such cutoff, valley and peak waves were classified according to frequency bands: delta (1 to 3.49 Hz), theta (3.50 to 7.99 Hz), alpha (8.00 to 14.99 Hz), beta-1 (15.00 to 24.99 Hz), beta-2 (25.00 to 34.99 Hz), and beta-3 (35.00 to 49.99 Hz).

As already mentioned, criteria were used for rejection of epochs of electroencephalographic signals containing artifacts; for this purpose every 10-sec epoch was divided into 2-sec periods. Appropriate criteria depend, of course, on the method of digitization, on knowledge of the characteristics of EEG's, and on the particular method of recording. The criterion used for muscle-spike rejection was 12 or more spikes within the 2-sec period, a spike being defined as a minima-to-maxima rise of at least 45 units within 16 msec. The extreme value criterion for rejection of a 2-sec period was a continuous period of at least 0.5 sec in which the value of the digitized signal was either greater than 337 or less than 174 units. These criteria were developed by trial and error with the advice of an electroencephalographer* and were verified by comparison with a standard reference (ref. 4). Examples of the artifact-rejection appear in figure 9.

As already mentioned, by failure to consider amplitude categories, some of the changes in the EEG signal are not brought out. This is particularly true for delta and theta waves at 8-Hz smoothing which show only very small changes throughout the entire recording. It was shown above that, when high-amplitude waves only are considered, there are marked increases in the number of delta waves during stages of deep sleep, and this will be demonstrated again in analysis-3.

The results for the bands having the most meaningful changes are shown in figure 10. The percentage of 2-sec periods that were rejected according to the criteria described above is also shown. The values are averages for 10-min periods.

The events marked on the chart and data-rejection graph will be discussed first. The following events were considered: pre- and post-lift-off periods, sleep periods (considered in detail below), eyes-closed periods, meals, housekeeping, and exercise. Some of these events are listed in a log supplied by NASA; others were interpreted by visual examination of the EEG chart by electroencephalographers* (ref. 3). The data-rejection graph shows no artifacts before lift-off, but a considerable amount of muscle-spike activity for several minutes just after lift-off. Figure 11 is a detailed graph of the percentage data-rejection for several minutes prior to and after lift-off; it shows no rejection prior to lift-off, peak rejection at 6 min after lift-off, and abrupt cessation of rejection after 23 min. Aside from lift-off, figure 10 also shows large amounts of rejection correlated with occurrence of meals and possible correlations with exercise and housekeeping. There is also a general trend of increasing rejection near 54:00 which would correspond to the increased loosening of the electrode.

The major changes in the frequency bands shown in figure 10 are associated with sleep or the eyes-closed condition. The log of the flight, supplied by NASA to Henry Ford Hospital, shows two periods of actual sleep by the command pilot: the second corresponds to the interpretation by visual examination, but the first does not. This lack of correspondence is also consistent with the computer findings.

Consider first the second sleep period. There are cyclic decreases in beta-1 (at 50-Hz smoothing cutoff) and beta-2 (at 50 Hz) which correspond to the deepest periods of sleep (as interpreted by the visual analyses). Both theta bands (50- and 25-Hz smoothing) show cyclic increases during the deep-sleep periods. Two effects seem to control the alpha activity: the alpha band, at 15-Hz smoothing, peaks only during the stage-0 or eyes-closed, resting state; but, at 50-Hz smoothing, increases are also seen during the

*Personal communication with W. R. McCrum, 1967.

*Personal communication with W. R. McCrum, 1967.

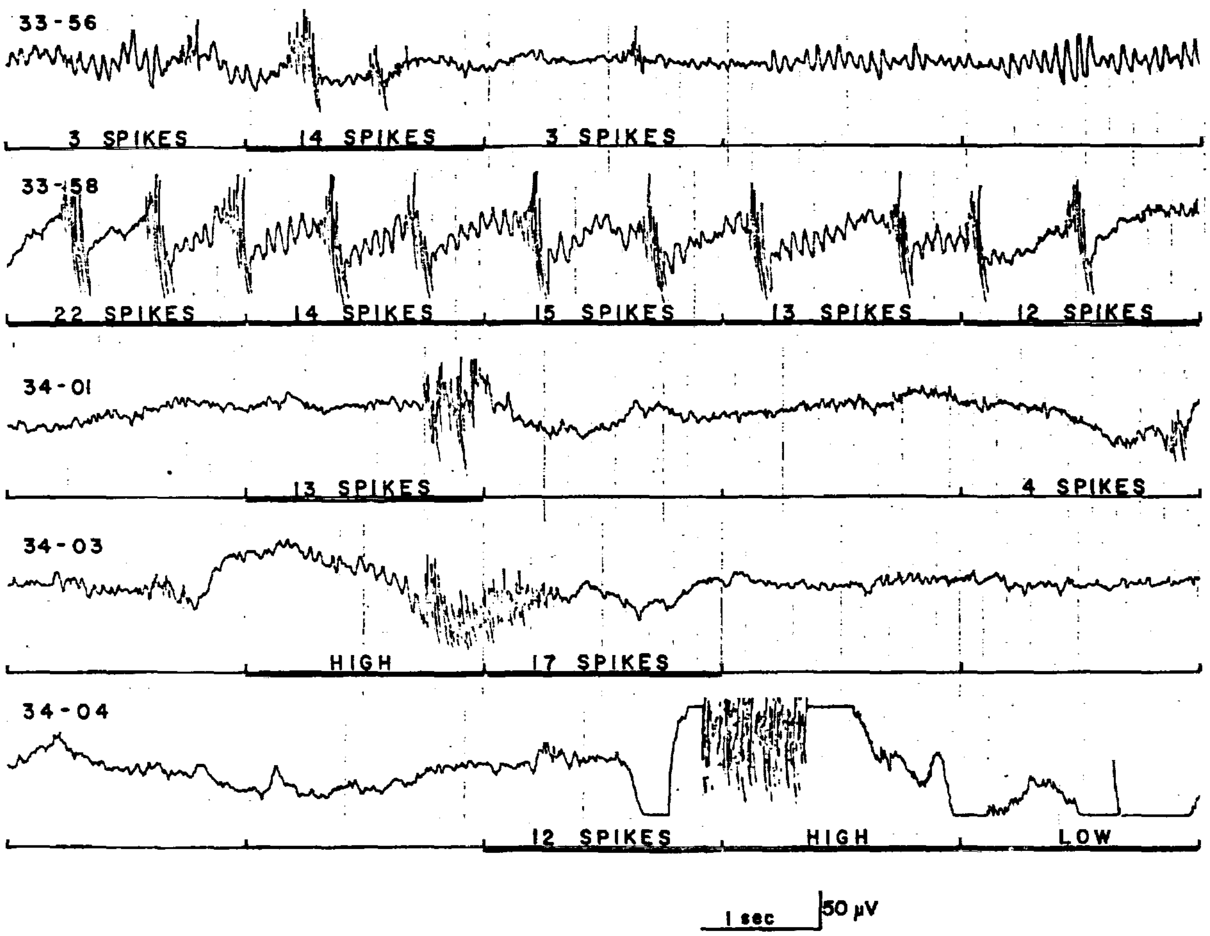


FIGURE 9.—Examples of artifact-rejection.

periods of deep sleep. These latter increases represent waves in the alpha frequency range which are superimposed on the delta and theta activity.

The changes just described are less apparent for the first sleep period, since there was a much smaller amount of deep sleep and its total duration was short. Small peaks are seen for theta (at 25 Hz); alpha (at 50 Hz) shows an increase during this sleep period; and there is a small decrease for beta-1. The sleep periods are discussed in much greater detail in the other analyses.

The eyes-closed period shows clear-cut increases for both alpha bands but decreases in theta (at 25 Hz) activity which is just opposite what was seen during the sleep periods. Beta-2 decreases somewhat during the eyes-closed period.

Apart from the sleep periods, most of the bands remain relatively constant throughout the flight; they are particularly constant during the first 7 hours. Only small differences are seen in the EEG when pre- and post-lift-off periods are compared (except in artifact-rejection); there is a small increase in beta-2 (at 50 Hz), and a small decrease in alpha (at 50 Hz) activity. In the 5-hour period just prior to the second sleep period some small trends in some of the bands seem to appear: all alpha bands tend to decrease gradually, while theta (at 50-Hz smoothing) and beta-2 tend to increase. One other finding is that the alpha band (at 15 Hz) seems to increase shortly after each meal.

Analysis-3—The program for analysis-3 is exactly the same as for analysis-2 except that

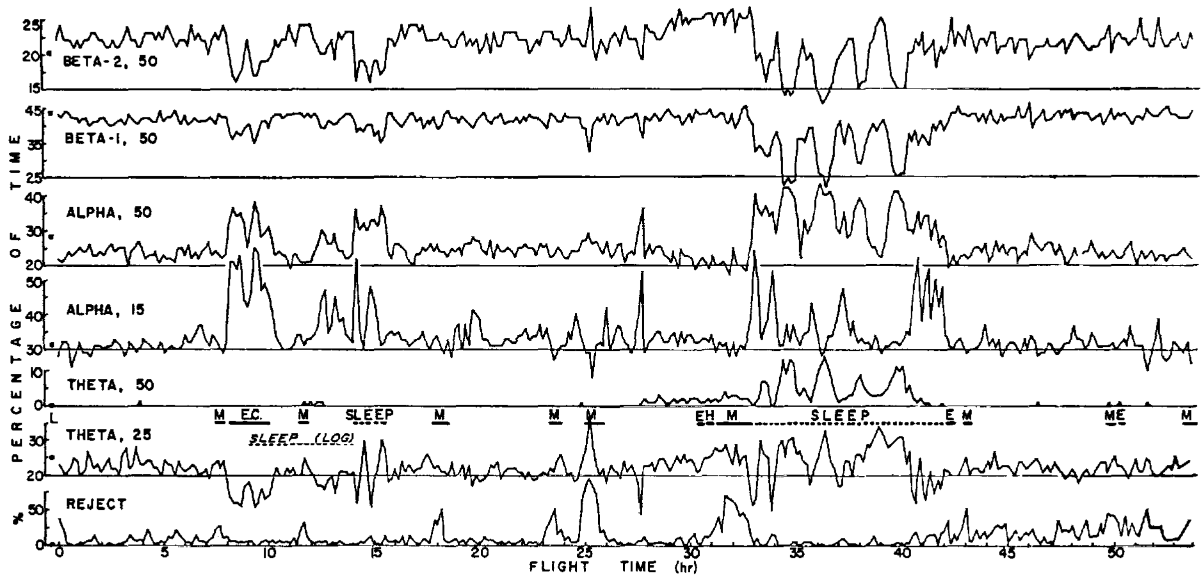


FIGURE 10.—Activity, in percentage of time, for the frequency bands shown.

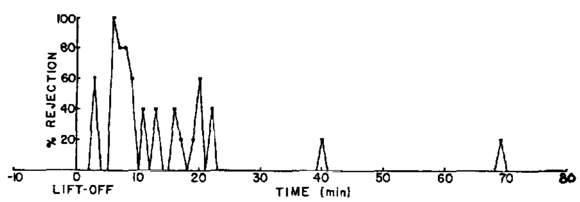


FIGURE 11.—Percentages of 2-sec periods rejected by the artifact-rejection criteria for the periods before and after lift-off.

waves with an amplitude below a certain criterion were not counted. With this criterion, changes in the EEG that were not apparent in the previous analysis are brought out quite clearly. The effect of the cutoff is rejection of the low-amplitude waves that might be considered "noise" or "statistical fluctuations" and essentially masking of the changes in higher-amplitude activity. The criterion used was linearly related to the wave duration and varied from 10 to 12 units for frequencies between 50 and 16 Hz and from 50 to 80 units for frequencies between 15 and 1 Hz. These criteria were found by experience to reveal most clearly the changes in EEG activity. This analysis was applied to the first sleep period (fig. 12) and part of the second sleep period (fig. 13).

Figures 12 and 13 show the percentage of time that the signal was in each of four frequency bands (recorded after different smoothing cutoffs): beta-1 at 50-Hz smoothing cutoff, alpha at 25 Hz, theta at 15 Hz, and delta at 8 Hz. The data represent means for 3-min periods (fig. 12) or 5-min periods (fig. 13), and the time of the beginning of each period is shown. At the top of figure 12 is a graph representing the visual analysis of the EEG performed by McCrum for the first sleep period. In figure 13 is a row of numbers representing the approximate average stages of sleep (ref. 3). It is apparent that the delta (at 8 Hz) and theta (at 15 Hz) bands are quite closely related to the depth of sleep. The alpha band (at 25 Hz) shows a peak in the drowsy period just before the onset of sleep. Some small increases in the alpha-frequency range, occurring at the peak of delta activity, represent faster activity superimposed on the slow sleep waves. The beta-1 band shows an almost perfect inverse relation with the depth of sleep.

To summarize, three versions of the basic SPC method were used for analysis of the data from flight. The first and relatively complex version provided complete categorization of the signal

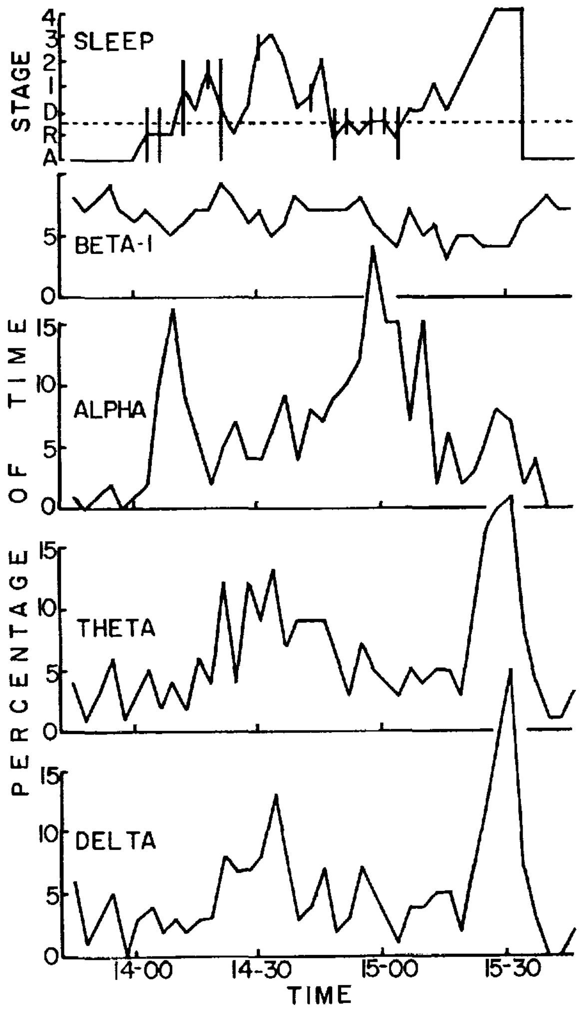


FIGURE 12.—Activities, in percentages of time, for bands (with amplitude cutoff): beta-1 at 50-Hz smoothing cutoff, alpha at 25 Hz, theta at 15 Hz, and delta at 8 Hz.

according to frequency, amplitude, symmetry, and smoothing cutoff. It was applied to only a few data and was used largely to explain the capabilities of the general method and as an aid in understanding of the results of the analyses that followed. The results of this analysis suggested which particular wave patterns characterized each stage of sleep, and the complexity for these waves was indicated by changes that resulted from successive digital smoothings.

The second and third versions were summary methods. In the second version the properties of amplitude and symmetry were not considered. This frequency analysis leads to some interesting findings in the higher frequency ranges but proves to be inadequate for showing changes in the delta region. The source of the difficulty is what might be called the "random character" or "noise" in the EEG signal. As successive smoothings are performed at lower and lower frequencies, large amounts of low-amplitude waves appearing in the delta range tend to mask changes in the high-amplitude categories. The third analysis showed that this masking could be easily eliminated by establishment of a cutoff such that only waves greater than a certain amplitude were counted. It was found that the high-amplitude delta waves followed very exactly the deep stages of sleep as interpreted by visual examination. Development of methods of data-summarization will be emphasized in future work on the SPC technique.

Briefly the following was learned about the EEG (channel-1) for the flight. There were only small changes in the data associated with lift-off. There was, however, a period of 23 min during lift-off in which large amounts of data were rejected according to the "muscle-spike criteria." Cyclic variations in depth of sleep were clearly shown by several of the measures used. The delta band (at 8-Hz smoothing cutoff), with a low-amplitude cutoff, followed very exactly the stages of sleep as interpreted by visual analysis. It was confirmed that the second sleep period followed the log (supplied by NASA to Henry Ford Hospital), but that the first sleep period occurred at a time indicated by the visual analyses and different from the time indicated by the log. At 08:00 there was a period of approximately 2 hours of strong alpha activity (eyes-closed, resting stage, and possibly light sleep). The criteria for data-rejection showed increased percentages of periods rejected at times corresponding to chewing and possibly to exercise and housekeeping. Increase in alpha activity appeared to follow shortly each chewing period. Near the end of the 54 hours of valid EEG recording there was a trend to increase in data-rejection that would correspond to gradual loosening of the electrode; it began at about 40:00. A small trend, that may have some importance, occurred over a period of about 5 hours prior to the second sleep period. All alpha bands tended to decrease gradually in

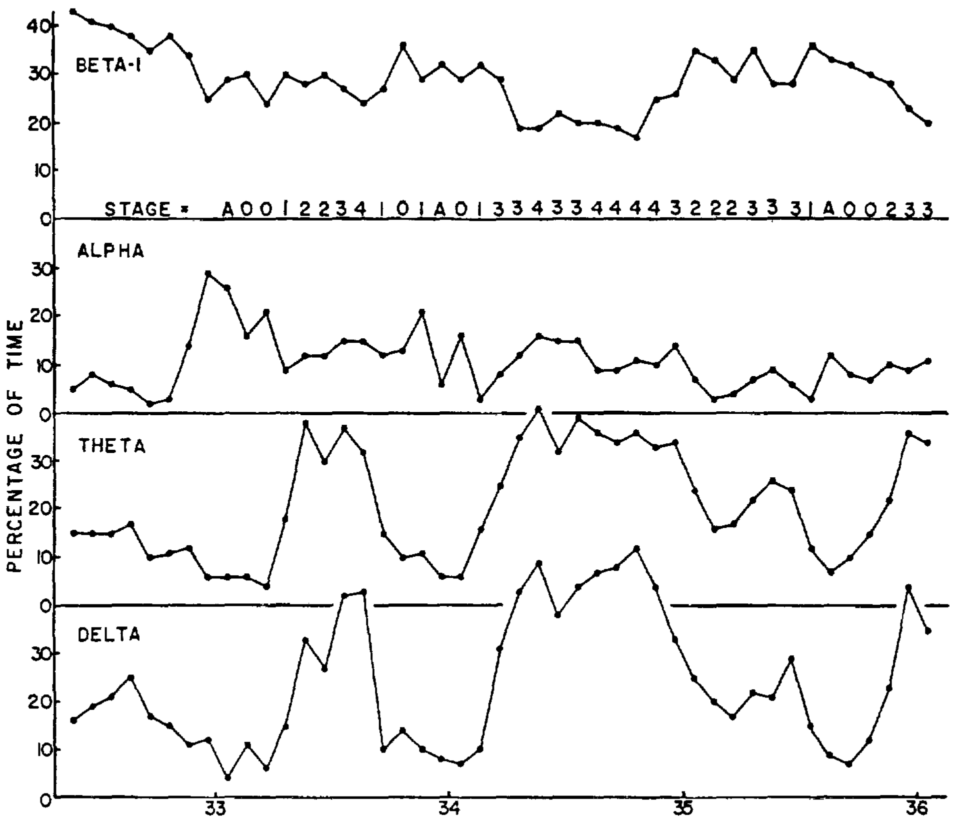


FIGURE 13.—Activities, in percentages of time, for each of four frequency bands (with amplitude cutoff): beta-1 at 50-Hz smoothing, alpha at 25 Hz, theta at 15 Hz, and delta at 8 Hz.

activity, while theta (at 50-Hz smoothing) and beta-2 tended to increase.

WEIBULL STATISTIC

Considerable work has been done in industry with a nonparametric statistical method using the Weibull distribution function (ref. 5). This method provides qualitative as well as quantitative evaluation of a distribution without requiring any assumptions about the distribution parameters. Experience has shown that the Weibull distribution more closely approximates physical and biological systems than does the normal distribution.

In consideration of any set of experimental observations or measurements as a population of individual statistics, this population can be defined as an ordered set and described by a cumulative distribution function having both form and magnitude. The form and magnitude of the total set from which the sample set was drawn are called the parameters of the total population, and these are best estimated from the characteristics of the cumulative distribution function of the sample. Thus any cumulative distribution function of a total set is composed of an infinite number of small distributions, each of which has its own unique parameters. For practical purposes these small distributions are generalized to one large distribution (the total sample) with one set of parameter estimates. If the members of this large population are essentially similar, it is said to be a unimodal distribution. If, on the other hand, there are two or more subpopulations that differ greatly in one or more parameters, the population is said to be bimodal or multimodal and is called a complex distribution. From experience it is known that no sample population is ideally unimodal. The variation of the sample from the ideal can be defined in two areas: (1) sample error due to too few samples failing to represent a large population, and (2) known or suspected competitive influences introduced into the sample.

For most conventional statistical procedures one must assume that the probability density function of the total population that is sampled is of the form

$$f(x)=[1/(\sigma\sqrt{2\pi})]e^{-1/2[(x-\mu)/\sigma]}$$

In such a function the mean, median, and mode coincide, and the shape of the function is such that the point of change of curvature lies σ-distance from the mean. It has been my experience (and that of others) that this function is rarely encountered in dynamic biological situations. Thus, conventional statistical procedures have little real meaning and furthermore can lead to erroneous conclusions.

Given the sample function, $P(X\leq x)=F(x)$, any distribution function can be determined from the equation

$$F(x)=1-e^{-[(x-\alpha)/(\theta-\alpha)]^{\beta}}$$

provided that $-[(x-\alpha)/(\theta-\alpha)]^{\beta}$ is nondiminishing and vanishes at some value of α. When α is assumed to be zero,

$$F(x)=1-e^{-(x/\theta)^{\beta}}$$

When $x=\theta$, the function evaluated at θ is

$$F(x)=1-e^{-1}$$

Thus we have a distribution function characterized by three parameters: (1) alpha, the minimum life; (2) beta, the slope or shape; and (3) theta, which is scalar. For most cases alpha can be assumed equal to zero and ignored.

Since differences in the mean and variance of two samples have no meaning unless the sample populations have the same shape, the first step in analysis is determination of the shape of the distributions. With use of the Weibull method, the first step then is evaluation of the slope beta. This was done by plotting of the ordered data points on Weibull Cumulative Distribution paper.

In the method presented, sample data points are plotted against median rank values, and a straight-line best fit is drawn to the plotted points. The median rank value is the value that has 50-percent probability of being larger or smaller than the true value and can be determined from the equation

$$\text{median rank}=(j-0.3)/(n+0.4)$$

where j is the jth-order statistic and n is the total sample. The cumulative distribution function

$$F(x)=1-e^{-(x/\theta)^{\beta}}$$

is rearranged as

$$1/[1-F(x)]=e^{-(x/\theta)^{\beta}}$$

Taking the $\log_e \log_e$ of both sides results in

$$\ln\ln\{1/[1-F(x)]\}=\beta\ln x-\beta\ln\theta$$

making the substitution

$$Y=\ln\ln\{1/[1-F(x)]\},\quad T=\ln x,\quad C=\beta\ln\theta$$

then $Y=\beta T+C$ which is linear plot in Y and T with slope β. Note that $1-F(x)$ is the cumulative probability of success. Weibull-probability paper is so constructed that the vertical scale represents

$$\ln\ln\{1/[1-F(x)]\}$$

but is graduated in terms of $F(x)$, the fraction falling within x. The horizontal scale is a logarithmic scale representing $\ln x$. A computer program was used that orders the sample data points, determines their median rank, and then computes the slope by a least-squares method.

The problem that arises in use of this computational method of determining the slope (beta) is that all distributions are then represented as simple or unimodal. My experience has shown that often this is not the case; rather the distributions tend to be complex. This fact can be determined by a graphic plot that connects each data point rather than a best-fit line; variations of this line from linearity then represent either sample errors or competitive modes.

It is my experience that in some cases the evaluation of beta is the only adequate information available from comparison of two sample populations; that is, the mean and variance of the two samples do not differ, but there is significant difference in the value of β. It has usually been found that this difference is due to introduction of a second mode into one sample that does not affect its mean and variance.

Quantitative differences between two sample populations can be determined by plotting of the 90-percent confidence bands about them. The method is the same as the one described, except that the 5- and 95-percent ranks are substituted for the median rank of the order statistic.

WEIBULL ANALYSIS OF FLIGHT EEG'S

The output of the Z/C technique provides a frequency count and normalized percentage times for a 15-sec epoch of EEG. Each 10 sec thus yields one sample point or value for the Weibull graph of a particular frequency band. After an arbitrary decision that a sample population of 60 points was sufficient to represent adequately the distribution curve, the computer was programmed to plot the Weibull distributions for each consecutive 15 min of analog EEG. On occasions during the flight recording, this 15-min length of record was disadvantageous; for example, some periods of sleep or a particular activity were less than 15 min in length. In such cases, any EEG information specific to that short period was diluted in the total 15-min sample. Shortening of the period of analysis would be costly in time and accuracy although the Weibull statistic can be quite powerful with a small sample of data. During the time before and after blast-off, the Weibull plots of 5-min periods of EEG were drawn by hand; this takes time and is quite prohibitive for large amounts of data.

One can also plot the 90-percent confidence bands about a Weibull distribution; that is, it is possible to plot the range within which 90 percent of all values of the distribution will lie. The plotting of these confidence bands has not yet been programmed for the computer, so they were used only in analysis of the pre- and post-lift-off data.

The command pilot's first sleep period was analyzed visually in a fashion resembling Maulsby's (ref. 3); we are in agreement to the minute for the onset and termination of sleep (fig. 14). The only disagreement was minor in the separation of stages-II and -III of sleep; we did agree on stage-IV. One should note that the Gemini 7 log supplied to Henry Ford Hospital appeared to err regarding the first actual sleep period, as it indicates that this occurred between 09:00 and

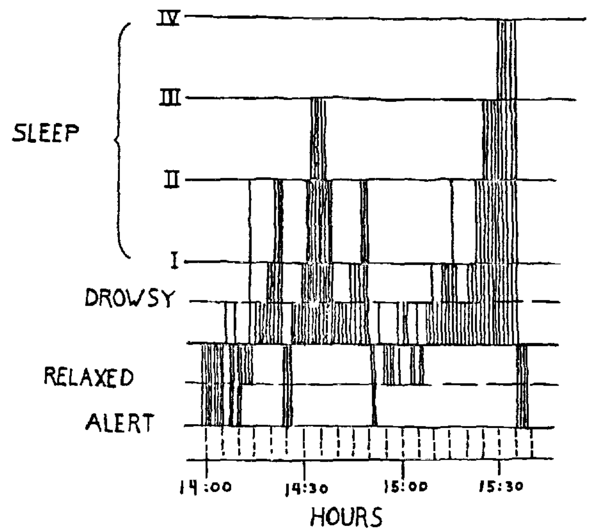


FIGURE 14.—Visual analysis of the first sleep period.

13:00 whereas our visual examination confirms Maulsby's conclusion (ref. 3) that this sleep took place between 14:21 and 15:35. The remainder of the record was perused visually for gross features such as chewing movements and presence or absence of sleep (without classification into stages), and also for long periods of alpha activity suggesting relaxation. An exact log of these observations was not supplied to Henry Ford Hospital.

The EEG taken during 10 min before blast-off was compared to that for 10 min following blast-off. Figure 15 shows the 90-percent confidence bands about the delta activity (0.5 to 3 Hz) for the states before and after blast-off. The fact that the two bands are widely separated indicates an extremely high probability that the marked reduction in delta activity after blast-off is a true observation. The abscissa is a log scale, and the medians of these distributions indicate that the percentage time of delta activity was halved after blast-off.

Figure 16 shows the 90-percent confidence band about beta-1 activity (15–25 Hz) after blast-off, along with the simple Weibull plots of beta-1 activity during the two consecutive 5-min periods just preceding blast-off. The indication is that the beta-1 activity almost doubled after blast-off, but the probability level is not quite as high as with the changes in delta activity. In the analysis

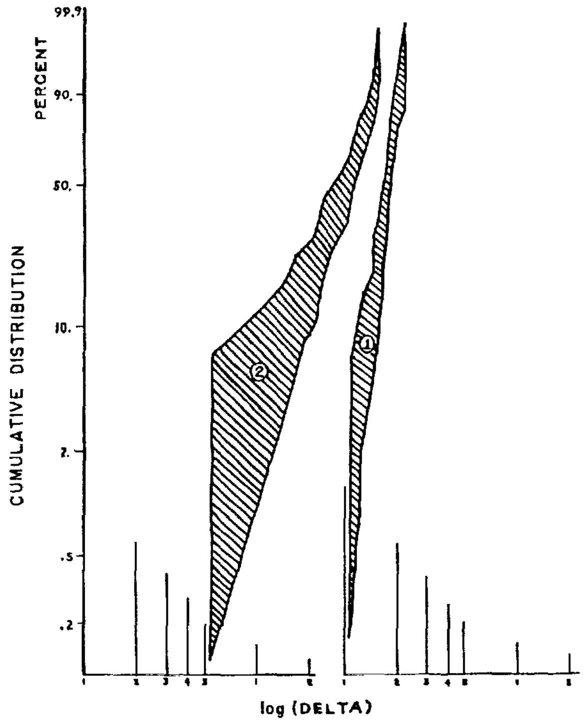


FIGURE 15.—Ninety-percent confidence bands about delta activity during 10-min periods preceding (1) and following lift-off (2).

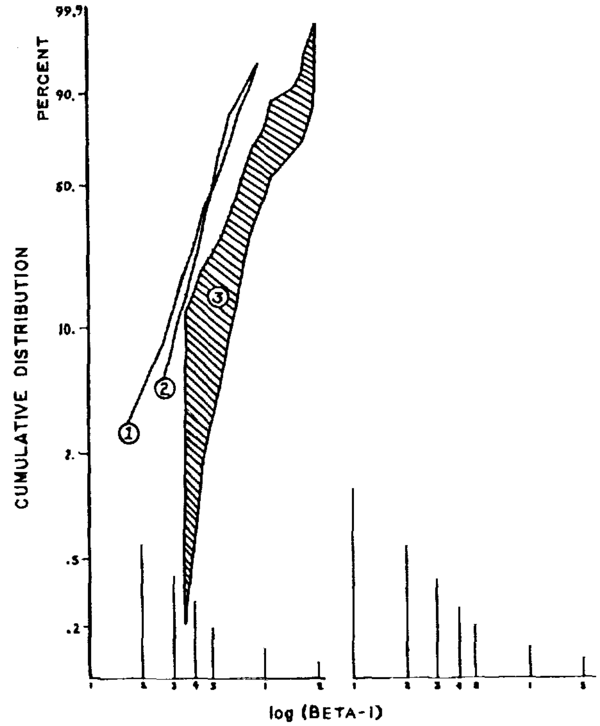


FIGURE 16.—(1) Weibull plot of beta activity from between 10 and 5 min before lift-off. (2) Beta activity between −5 min and lift-off. (3) Ninety-percent confidence band about the beta activity for the 5 min immediately following lift-off.

of alpha activity (fig. 17) the 90-percent confidence band is for the activity (8 to 15 Hz) following blast-off. Only the simple distribution of the pre-blast-off theta is drawn. Only in the midrange of each sample is there significant difference in the amount of alpha, with the post-blast-off record showing the greater amount; the tails of the samples are not different. This fact may suggest a change in the state of physiology of the brain that is greater than would be reflected by simple increase in the mean amount of alpha.

Analysis of theta activity yielded no evidence of change between the periods before and after blast-off. The Weibull plots of the two sample periods were almost identical.

Weibull distribution plots were made for each consecutive 15 min of EEG for the entire 53.5 hours. There were separate graphs of each frequency band: delta, theta, alpha, and beta-1. This made a total of 856 individual graphs of data derived from the Z/C technique. A similar set of 856 graphs has been derived from the SPC technique for processing of the analog EEG. In this report only the plots derived from the Z/C program have been reviewed in their entirety, and only a few chosen graphs derived from the SPC method have been reviewed.

My first attempt to compare the Weibull plots of a given frequency band was by simple superimposition of one graph over the other. When this was done the range of alpha activity (that is, the variation in quantity) was quite small over the entire orbit; likewise the general slope or shape parameter was similar. This procedure did yield some ideas for future analysis that will be discussed later, but for our present purpose it appeared impractical.

Attention was then limited to the sleep periods. The first sleep period lasted about 1.5 hours and was interrupted frequently by periods of arousal. Each of the 15-min periods of EEG of this sleep

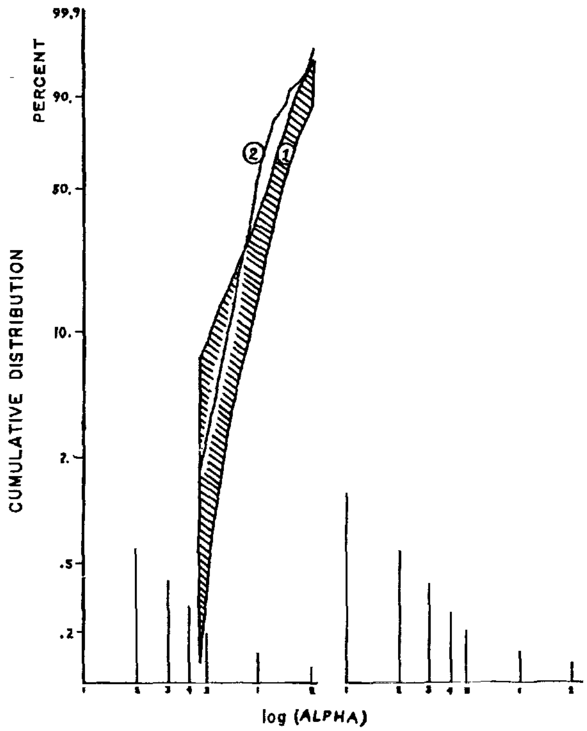


FIGURE 17.—(1) Ninety-percent confidence band about the alpha activity for the 10 min following lift-off. (2) The simple Weibull plot of the alpha activity for the 10 min preceding lift-off.

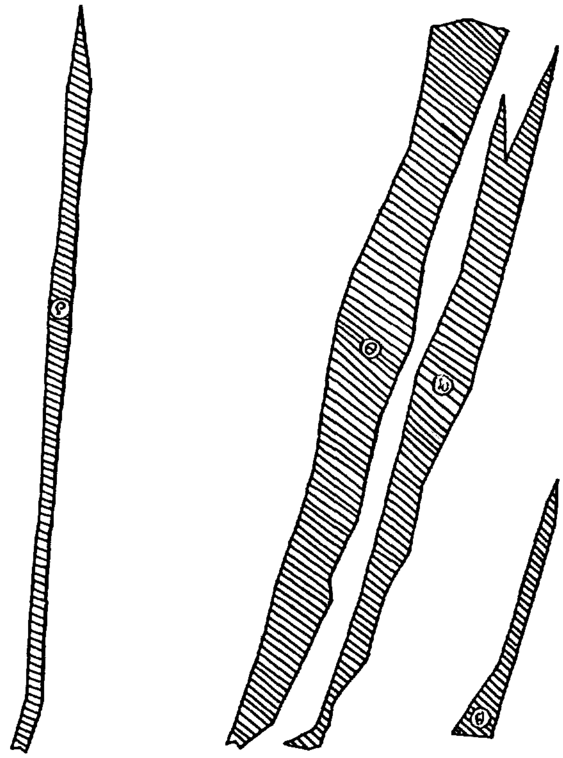


FIGURE 18.—Range of Weibull plots of stage-IV sleep for the frequency bands delta, theta, alpha, and beta-1; data derived from the Z/C technique.

episode was a mixture of sleep and arousal, and the collection of Weibull plots over this 1.5-hour period showed no features that would permit labeling of any particular plot as representative of sleep.

The second sleep period covered a period of about 8 hours. The Weibull plots of the consecutive 15-min EEG samples over these 8 hours were examined collectively by superimposing one over another. Primarily on the basis of the plots of delta activity, four 15-min epochs of EEG were separated from the total sleep group. When these four periods were related to the visual analysis of the EEG, they coincided identically with the periods of stage-IV sleep that were of 10-min duration or longer.

The ranges of the Weibull plots of delta, theta, alpha, and beta are depicted in figure 18. There is complete separation of each band from the other, even though the ranges of alpha and theta are quite wide. When the data from the SPC method were used, the ranges of alpha and theta activities narrowed considerably, but, on the other hand, the total quantity of each increased manyfold. This observation poses some questions about which methods of analog processing should be used; each seems to have certain advantages over the other.

The various stages of sleep other than stage-IV could not be evaluated by this means of comparing their Weibull plots. One reason for this was that in any 10-min period, represented by a single graph, several stages of sleep were represented; thus each Weibull plot represented a mixture of the different stages of sleep.

One other finding needs some clarification: It must be remembered that the Weibull plot yields a shape parameter that is some measure of the operating characteristic of the system under study; in this case it is the physiology of the brain underlying electroencephalographic activity. This shape parameter may change considerably with or without alteration of the parametric mean values of EEG analysis. When this happens, it may be assumed that the physiology has changed in some way that is not reflected by a simple quantitative measurement of an EEG frequency band.

When the slope or shape parameter of the alpha activity, obtained from the SPC method, was studied in detail, it became apparent that the various graphs fell into groups. All these graphs represented compound or multimodal distributions; that is, a single Weibull plot was not a simple straight line but a series of straight lines at various angles to one another. When the first 30 hours of the orbital flight was examined in this fashion, it was apparent that the 120 graphs representing this time fell into eight major groups. The important observation was that all the graphs for the first 3 hours of the flight belonged to just one group; thereafter there was complete intermingling of the members of the other groups in a seemingly random fashion. The median amount of alpha activity was similar throughout the entire 30-hour period. Figure 19 shows the characteristic plot of the first 3 hours after blast-off, along with some examples of the other shapes found later.

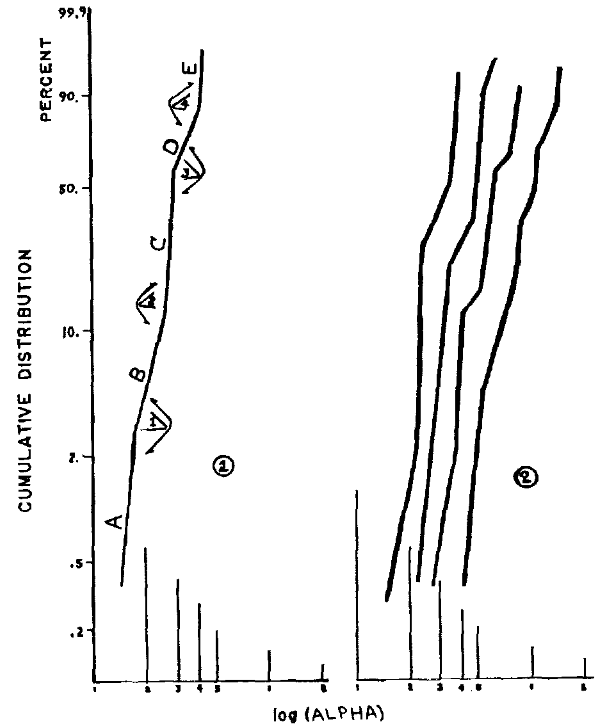


FIGURE 19.—Weibull plots of alpha activity, derived from the SPC technique.

In summing-up the Weibull analysis it can be said that if significant quantitative differences in the EEG are present they can be determined to be significant with only a small sample of the EEG. Examination of the periods immediately before and after blast-off bears this out. The periods of stage-IV sleep could also be determined quite accurately. Examination in detail of the slopes of the individual graphs, particularly alpha activity, could be the most important finding of all; unfortunately we still have no satisfactory method of examining and classifying the hundreds of slopes that are gathered from a long period of EEG analysis.

Several problems are involved: For one thing the exact distribution functions for most of the sample periods are unknown; none of them is normally distributed or even symmetrically distributed. We have accepted use of the Weibull distribution as a close approximation. In many cases this seems true, but in many others it is obviously not; thus the graphic analysis of distribution functions must be investigated. If different kinds of distribution functions can be shown to exist in different EEG samples, these in turn can be directly related to changes in physiology and/or behavior. When this correct graphing of distribution functions has been accomplished, computer pattern-recognition programs must be developed to define the different graphs as finitely as possible. This poses further problems. Since, on the basis of present observations, quantitative differences in the EEG appear significant when there are marked changes from alertness to deep sleep, the analysis of fine changes in behavior (both affective and physiological) should be directed toward the tails of the distributions where behavioral changes are most reflected. A proper pattern-recognition program or technique would then require some knowledge of the "statistics of extremes."

These are only some of the problems in development of the "nonparametric" approach to EEG analysis. My findings in this study, using the Weibull statistic as a first approximation, seem to justify further efforts along this line.

GENERAL DISCUSSION

I have reported my analyses of the flight EEG records by routine visual EEG analysis, by parametric analysis of data by the zero-crossings technique, and by a pattern analysis called smoothing and peak-counting. Finally we have applied Weibull statistic to the data derived from both the Z/C and the SPC techniques. A routine visual analysis was made of the two major sleep periods, the first appearing at about 14:00 and lasting roughly 1.5 hours and the second appearing about 33:00 and lasting approximately 8.5 hours.

The parametric and SPC analyses revealed another period of light sleep between 08:00 and 10:00. Visual reevaluation of this period of the EEG did reveal the presence of stage-I sleep throughout and a few minutes of stage-II sleep near the end of the period (fig. 20) when the criteria of Dement and Kleitman (ref. 2) were strictly applied. We would expect the periods of lighter sleep to be missed on routine visual analysis when not followed by periods of deeper sleep that serve to cue their presence for the clinical electroencephalographer. This fact demonstrates that our described computer techniques can recognize the various stages of sleep independently of one another.

On review of the results of our various methods of analysis of the EEG it appears that quantitative analysis by use of computer techniques is superior to routine visual analysis. In particular there are two major advantages: first, the computer provides a consistent numerical analysis, and second, the results are not dependent on subjective inference. The ideal method of analysis by computer obviously has not been attained, but we believe that through a number of techniques we have effected a useful analysis. Of these methods the one to be used in a particular case will be determined by such factors as quality of the raw data, cost, and time for the analysis. In upgrading of the computer techniques a universal system will, we trust, be found that incorporates the best features of the various computer programs.

Each of the techniques reported here has its own individual merits. The Z/C technique provides data that, when summarized with standard parametric and nonparametric statistical techniques, reliably indicate major changes in behavioral states that are currently definable and now receiving significant scientific attention. The Z/C technique is the most rapid computer technique operating with acceptable real time; consequently it is the lowest in cost since it also lends itself to a variety of small, compact, inexpensive computers.

The SPC technique is particularly capable of handling noise in EEG signals by rejection criteria and by not being bound to a base line; it provides detailed categorization of the signal

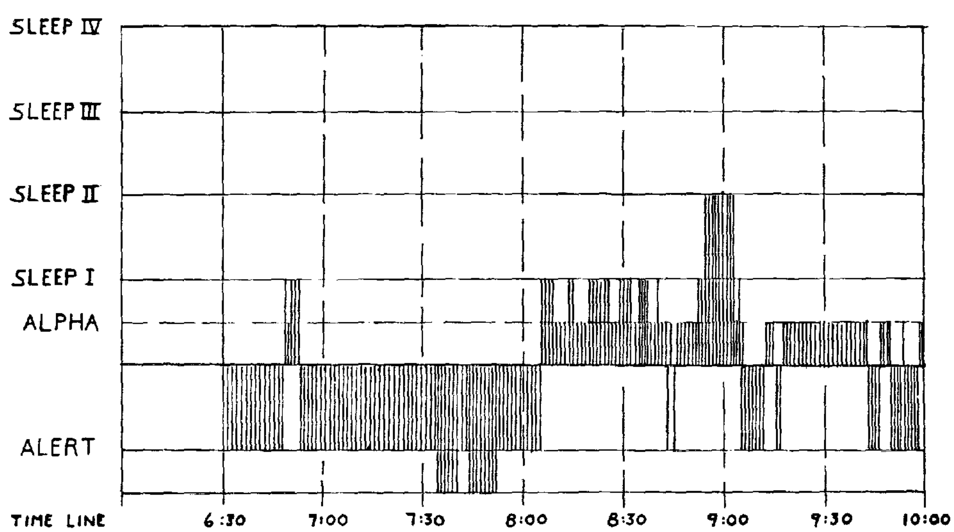


FIGURE 20.—Visual analysis of EEG during "eyes closed" (sleep?).

according to frequency, amplitude, symmetry, and complexity (superimposition of faster or slower waves). This technique is also a relatively rapid computer program and could be used on-line.

The Weibull statistic is a useful research tool to assist the modeling of the neurophysiology underlying the EEG. It served the purpose in this report of emphasizing the complex variability in data obtained by Z/C and SPC techniques. For on-line analysis of sleep, the Weibull statistic offers no advantages.

In my opinion improvement of the electrode system and substitution of a noncontinuous program for the continuous EEG recording would prove more efficient and less inconvenient to the astronauts.

ACKNOWLEDGMENT

The following members of the staff of the Department of Behavioral and Neurological Sciences, Henry Ford Hospital, contributed to this work: W. R. McCrum, Ph.D., T. E. LeVere, Ph.D., R. M. Lee, Ph.D., and H. van den Ende, M.S.E.E.

REFERENCES

1. Bartlett, M. S.: Square-Root Transformation in Analysis of Variance. J. Statist. Soc. Suppl., vol. 3, 1936, pp. 68–78.
2. Dement, W.; and Kleitman, N.: Cyclic Variations in EEG during Sleep and Their Relation to Eye Movements, Body Motility and Dreaming. Electroencephalog. Clin. Neurophysiol., vol. 9, 1957, pp. 673–690.
3. Maulsby, R. L.: Electroencephalogram during Orbital Flight. Aerospace Med., vol. 37, 1966, pp. 1022–1026.
4. Gibbs, F. A.; and Gibbs, E. L.: Atlas of Electroencephalography. Vol. 1. Addison-Wesley Press, Inc., 1950.
5. Weibull, W.: A Statistical Distribution Function of Wide Applicability. J. Appl. Mech., vol. 18, 1951, pp. 293–297.
6. Johnson, L. G.: Unpublished notes and lectures.

CHAPTER 10

TRACKING OF AN ASTRONAUT'S STATE BY PHYSICAL MEASUREMENTS OF SPEECH

Louis V. Surgent

SUMMARY

This study is addressed to the question of whether it is feasible to use measurable speech parameters for detection and tracking of the changes in state of an astronaut. Techniques for assessment of changes of state are developed to serve as criteria. The most detailed of these, Probable-state analysis, uses the "stream of behavior" concepts of Barker (refs. 1–3) to organize situational and behavioral data for raters including speech communications. Procedures for rating the state of the pilot during each communique, on a set of 10 situationally and behaviorally defined "probable states," are specified.

Automatic speech-processing techniques were found inapplicable to large sections of the onboard recordings supplied for analysis. Instead, oscillographic measuring techniques were devised to simulate a feasible automatic speech-processing system and applied to about 68 min (5121 syllables) of astronauts' speech. Communiques were segmented into periods of uninterrupted speech, called Pause Groups or Groups, by a criterion silence (170 msec) selected to distinguish between "articulatory" and pausal silences.

Of the new measures developed, Group Highest Pitch and a ratio formed by dividing the duration of the Group Last Syllable by the average duration of the remaining syllables, named the DURLL Ratio from its FORTRAN designation, are most promising. The standard deviation of algebraic, first-order, serial differences in successive syllable rates—computed as the reciprocal of each syllable duration—showed responsiveness to situational changes, especially when normalized in a ratio with the standard deviation computed without regard to sequential dependencies.

Syllable Peak Amplitude and Syllable Rate (computed as above and exclusive of pausal silences) also showed evidence of useful validity. Group Duration and Syllables per Group were weaker but clearly responsive. Pausal silences within communiques were distributed poorly and showed no evidence of validity.

Methodological considerations are emphasized, and suggestions for research, development, and application are offered.

INTRODUCTION

Detection and tracking of the changes of state of an astronaut from the physical parameters of his speech require discovery of measures that covary with state, and development of weighting functions that transform sets of these measurements into identifications of states and state-intensity predictions. This is a large order, with many unresolved problems regarding criteria (state) and predictor (speech). Substantial progress toward meeting these requirements, rather than complete fulfillment of them, was therefore the aim of this small, 1-year contract.

Within this frame, initial attention was given to criterion questions such as the following: How should the state of an astronaut over a period of time be specified? What set of labels should be used? What evidence will warrant assignment of one label rather than another—or of two or more labels simultaneously? How should variations in the degree (level or intensity) of a particular state be designated? Or should we settle for well-established physiological measures, such as heart rate, and attempt to predict them?

Since the validation methods eventually used were considerably less sophisticated than those planned, due to funding limitations, we proceed directly to the measurement of speech parameters and to the results from the simpler methods. The criterion problem is discussed more basically in two appendices.

MEASUREMENT OF SPEECH PARAMETERS

Possibility of Wholly Automatic Analysis of Speech

To be acceptable, speech monitoring must be inexpensive and convenient as well as valid, hence, the requirement for investigation of wholly automatic processing despite the problems foreseen. The main hope in view of capsule-power limitations lies in development of a spacecraft vocoder. The transmission of speech measurements digitized on board would provide intelligible speech with minimum bandwidth and power and would facilitate computer processing of selected parameters for pilot monitoring.

The current use of limiting circuits to increase transmitter-power utilization results in "speech clipping," with consequent gross spectral distortion as amplitude enters their region of operation. Thus acoustical analysis is complicated. In addition, the weight-comfort relation in continuously worn headsets dictated the use of microphones that attenuated drastically frequencies in the region of the voice fundamentals; their spectral envelopes differed, and some appeared marginal.

The on-board recordings supplied for this study reflected these hardware limitations; they also contained extraneous capsule noises, tone signals, and ground communiques, as well as receiver and other electrical noise, especially in intervals between astronauts' speech. On the other hand, the amplitude variability, spectral distortions, and intense random noise often introduced by atmospheric propagation were absent.

The types of error made by automatic equipment in digitizing such speech can be specified quite well without laboratory trial. However, speech-quality criteria normally applied to the synthesized output of such equipment in bandwidth-reduction studies are replaced here by the much less stringent and lesser-known requirements of probabilistic state-predictions.

Efforts to assess automatic processing for this application and with these recordings led to the following conclusions. First, almost every time function examined, including the second formant, contained enough information to justify its inclusion in a set of functions displayed oscillographically for human discrimination and measurement. Second, a sizable computer-programming effort, preceded by a substantial amount of engineering, was required to get anything much from automatic analysis of these recordings. Only the rectified and smoothed amplitude time function was sufficiently reliable in specifiable regions—over syllable nuclei, for example—to provide a start in this direction. The third conclusion is that automatic spectral analyses by Starkweather (ref. 4) in his clinical investigations, and the related techniques of Popov (ref. 5) in his study of two Russian astronauts and 15 actors rendering, "This is Diamond, I read you," etc., would be degraded seriously by these frequency-domain distortions and sounds with nonrandom components.

Semiautomatic Approximation to a Realizable Automatic System

There was good reason to continue the study despite temporary abandonment of wholly automatic processing and substitution of oscillographic techniques. Apart from the scientific value of relations that may be discovered, a semiautomatic system can be developed for digitizing of samples of speech, either as occasional checks (in the way in which blood-pressure readings are now taken) or distributed over time to allow for processing. This could hasten complete automation if the search for valid measures is confined to variables digitizable in a spaceship or moon-orbiting laboratory either with minimum additional circuitry and data load or as part of an on-board vocoder.

The oscillographic analysis reported here was guided by these objectives. They do not in themselves represent feasible operational procedures but are aimed at discovery of applicable measurements and relations.

Performance and Digital Encoding of the Basic Acoustic Measurements

Specimen oscillographic record—Figure 1 is an oscillographic excerpt of a pilot's communique beginning: "Main chute on green. Chute is out in reef condition at 10 800 feet and beautiful chute. Chute looks good. On O_2 . . . ." The acoustic

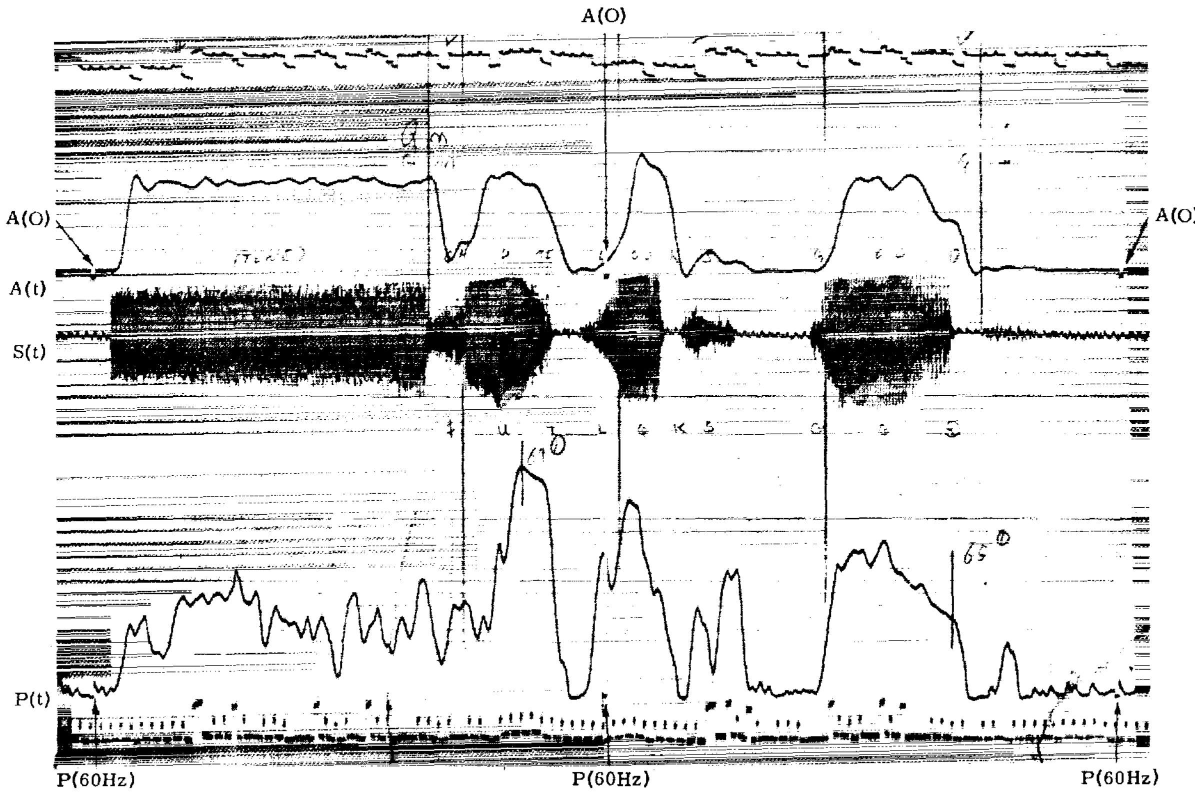


FIGURE 1.—Oscillogram of pilot saying "Chute looks good."

energy preceding the speech is thought to be a set of beat frequencies generated by two transmitters, one aboard a rescue plane and the other apparently aboard a rescue vessel. The recorded suppression of this energy, as the pilot begins his transmission, typifies the operation of VOX and push-to-talk circuits on noise and extraneous signals processed by the receiver.

The "speech pressure wave form" or "speech time function," $S(t)$, gives the clearest indications of the onset, termination, and nature of major speech events. Vertical displacements of from -0.6 to 0.6 in. were set equal to a 3.0-V peak-to-peak calibration tone recorded on each channel.

The "amplitude time function," $A(t)$, is a rectified and smoothed transformation of $S(t)$ with 1.5 in. equal to 1.28 V rms. Measurements are made from zero reference marks, $A(0)$, and corrected by the amount of calibrated attenuation or amplification necessary to bring speech amplitudes, over a flight phase or other period, within the dynamic range of recorders and other equipment.

The voice's fundamental frequency, attenuated by microphones, was reconstructed by preprocessing of circuits of the full-wave-rectification type and fed to a vocoder pitch-tracker to produce the "pitch time function," $P(t)$. Pitch-tracker outputs of zero and 5.0 V were equated respectively to a 60-Hz reference, P(60 Hz), and a 300-Hz tone. Input amplifiers to the oscillograph were then adjusted to place 250 Hz 1.52 in. above the P(60 Hz) reference marks which are recorded once each second. An empirically derived calibration equation was used to convert measured displacements over a wide range to cycles per second. A BCD serial code, identifying hours, minutes, and seconds, is along the bottom. A synchronous 10-pulse-per-second line along the top is used with an interpolating template to measure event times to centiseconds.

Pause Groups as response units—From a behavioral standpoint, examination of the $S(t)$ indicates the importance of distinguishing between what we called articulatory silence and pausal silence. The articulators are highly constrained with respect to position, movement, and timing during articulatory silences—e.g., between the "t" of "chute" and the "l" of "looks," before the release of the "k" of "looks," and between the "s" of "looks" and the release of the "g" of "good" (fig. 1). During pausal silences, on the other hand, the articulators are free of linguistic constraints. Inhalations occur only during pausal silences.

A criterion pause was then sought that was long enough to avoid segmentation during articulatory silences and short enough to assure segmentation by breath and hesitation pauses. A value of 170 msec was found to work well for the quite rapid speech of pilots in the tapes analyzed; this was detected as a 0.7-in. extent on the oscillographic records (4.25 in./sec). The duration of all pauses of 170 msec or longer was recorded with a view to investigation of the function relating the optimum criterion pause (inversely) to speech rate, but time and other constraints did not permit this ancillary analysis. Periods of speech bounded by criterion silences were called Pause Groups or simply Groups, being analogous to the breath groups or segments bounded by inhalations. On oscillographic readouts, Q's and G's, with "on" and "off" affixes, designate onsets and terminations of communiques and Groups.

The concept of segmenting of speech by pause durations for behavioral studies was proposed (refs. 6 to 8) to eliminate the human judgments required by Chapple's interaction chronograph. Verzeano investigated distributions (Poisson) of segment lengths as a function of criterion pauses of from 100 to 900 msec. Distributions generated by criteria of 400 msec or more showed irregularities which he attributed to "respiratory rhythmicity"; articulatory and pausal silences were not distinguished.

Brady's (ref. 9) concern with improvement of techniques for measuring speech level led him to analyze periods of measurable speech energy and silence in staged telephone conversations. He distinguished "intersyllabic gaps" from "listener-detected pauses" and found that 200 msec form a boundary between the two; he called the speech segments marked by such pauses "spurts." Brady's terminology would have been adopted except that we read his report some time after submission of the project report.

The term "articulation pauses" was used once by Goldman-Eisler (ref. 10) in a study of pause distributions; she selected 250 msec as the boundary separating these from hesitation and breath pauses. She also worked extensively with breath groups, and some of the measurements and indices that she developed from them (refs. 11 to 13) have analogs on Pause Groups.

Pause Groups provide an excellent starting point in analysis of communiques for information about the state of a speaker. First, the basis of segmentation is objective, as Verzeano emphasized, although one cannot gloss over the problems of discrimination of low-level speech energy from "silence" in real-world systems containing noise. In a wholly automated system, improvements in segmentation will be attainable by addition of other, more-complex, physical criteria. Second, Pause Groups exist in every language and can be identified without translation. Third, Pause Groups appear to have some linguistic and physiological relevance. Henderson, Goldman-Eisler, and Skarbek (refs. 14 and 15) provide the most interesting data. In reading, 100 percent of inhalations occurred at grammatically appropriate junctures, compared with 69 percent in spontaneous speech (cartoon stories). In reading, inhalations occurred in 77 percent of the pauses, compared with 34 percent in spontaneous speech.

The spontaneous-speech samples were further analyzed on an x-y plotter; the pen moved vertically during silences of 100 msec or longer and horizontally during speech or shorter silences. Alternating periods of hesitant and fluent speech were indicated by cyclic changes in slope, with speech during periods of steeper slope showing a larger number of "Ah's," false starts, and inhalations at ungrammatical junctures. These were interpreted as periods of "planning," which facilitate the fluent speech in the second phase of each cycle. No attempt was made to establish that behavioral processes ordinarily called planning really occurred. The possibility that the higher incidence of hesitation phenomena is associated with the greater information content of the thematic decisions, required at certain points in the developing cartoon story, is not mentioned.

However, it has been noted that it is more compatible with earlier demonstrations (ref. 16) that the length of a hesitation pause between two consecutive words varies inversely with the transitional probability. It may well be that the principal difference between the two processes is that the transitions noted here are between "kernels," phrases or larger units, instead of between words.

Goldman-Eisler (refs. 10 and 17) also found hesitation behavior somewhat less evident when subjects described a cartoon than when they attempted to summarize concisely the main point. With practice (seven repetitions with the same cartoon), pauses were substantially reduced and speech rate increased.

In addition to indicating the complexity of pause-related phenomena, these data place in perspective Fonagy and Magdics's (ref. 18) finding that over 95 percent of inhalations occurred at sentence or phrase boundaries, since their analysis was based mostly on reading and highly formalized speech such as sports broadcasts. They also illuminate our finding—in a spot check of an astronaut's speech during a launch—that about 85 percent (30 of 35) of pauses longer than 170 msec occurred at phrase boundaries as identified by a linguist before application of the 170-msec pause criterion. The many simulation runs performed by astronauts, before flight, undoubtedly resulted in establishment of speech responses and formats under control of the many fixed sequences of stimuli presented. Another large share of astronauts' responses are "descriptive" and may be expected to exhibit properties more closely resembling cartoon description than summary of the main point.

The fourth point is that the Pause Group can serve as the "experimental unit" or "unit of analysis." This applies in the sense that ideally all speech measurements and speaker-state assessments would be assembled on a Pause Group basis (see Appendix A).

Syllabic segmentation of Pause Groups—The simplest method of automatic marking of syllable onsets consists in low-pass filtering, rectification and smoothing of the speech pressure wave form, and setting of a threshold on the first derivative of the output. These circuits have also been used to measure speech rate (ref. 19). More complex, automatic, segmentation techniques are being developed at General Dynamics and elsewhere to avoid the susceptibility of these circuits to "misses" in certain phonemic environments and with rapid or slurred speech, and to "false detections" due mainly to the extreme sensitivity of the first derivative to noise.

The simpler circuits were chosen as the model to be simulated because they can be set to approximate perceived syllable onsets (ref. 20). The linguist who scored the oscillographs used three different rates of onset, equivalent to three values of the first derivative, as "anchoring points" in his judgments and placed marks to approximate perceived onsets. Additional guidelines were specified for marking of instrumentally troublesome onsets in the presence of semivowels, including the intervocalic "r", voice fricatives, and nasals.

Encoded data—The seven-digit format needed for encoding of Irig-B times to centiseconds was used for all event times, measurements, and other data. A two-digit function code was added to identify the decoding rule. A state diagram defining the structure and permissible sequences of function codes was drawn up to facilitate encoding and programming. The event times encoded were the onset and termination of each communique and of each Pause Group within it, and syllable onsets. "Syllable peak amplitude" (ARMS)* was an obvious choice, being an easy measurement, with some data relating it to the state of the speaker (ref. 21).

The literature at the time provided no guidance for scoring of the noisy output of a pitch-tracker. The discriminations necessary for the Fairbanks-Pronovost (ref. 22) measures, for example, could not be made. Intensive study of this function led to the conclusion that its maximum excursion over a Pause Group—exclusive of second harmonic seizures, which are easily recognized—was the only theoretically satisfactory, reliable, and potentially instrumentable measurement available on every Group; this was named Group Highest Pitch (PHH). To check the reliability of this measure and to provide estimates of its midvalue for each Group and of its variability

*Names (of variables) assigned for computer programming are used throughout; A (amplitude) was measured in volts rms.

over Groups, the second-highest pitch and the lowest and second-lowest "readable" pitches were also encoded. Only the PHH data were analyzed.

These preliminary assessments were verified not only by data from astronauts, as will be seen, but also by a subsequent report (ref. 23) in which grammatical sentences, uttered with simulated emotions or culled from classroom lectures and discussions, were rated for emotional content by listeners. The maximum pitch for each sentence correlated well with certain ratings. Correlations were somewhat higher with simulated emotions than for the 27 sentences (each repeated at least once, making 60 utterances) of the classroom instructor.

Other information included routine identifications, amplitude-reference and calibrated adjustments, and provision for state-relevant data. A simulation of the "voice-unvoice-silence" segmentation of a typical vocoder, together with an alphanumeric code identifying each phoneme, was tried and found feasible but abandoned for speed in the processing.

The measurement setup—All measurements were made and encoded manually since an oscillograph digitizer was not then available to us. Feed and take-up reels were secured 45 in. apart with the oscillograph paper supported by a metal bridge. A 32-in., clear plastic strip, with a machined straightedge, was aligned with the pitch-reference marks, P(60 Hz), and clamped. A specially designed measuring template rode along this strip, permitting time, amplitude, and pitch readings over about 30 in. (7+ sec) of speech in one setting. Oscillographs were "permanized" so that measurement was possible under ambient illumination.

The quantity of speech analyzed—Twenty 100-ft rolls of oscillograph paper, recorded at 4.25 in./sec, were completely processed for computer analysis; they contained 497 Groups and 5121 syllables, spanning more than 64 min of flight and the last two min of two countdowns. The exact distribution over flights is not listed here, to avoid explicit pilot-identification.

RESULTS

A number of strong and interesting relations were found between the speech measurements and emotionally toned flight events, despite the incompleteness and simplicity (due to funding limitations) of computer data-reductions, flight-phase comparisons, and graphic analyses.

Group Highest Pitch

Pitch variations during countdown, launch, and initial weightlessness—The joyful anticipation expressed in "Boy, can you imagine, here we go," just 2 min before lift-off (fig. 2), may have been suppressed somewhat by the physical discomforts of a long countdown and the tense readiness of the last moments. Relief from these, and the joy of "We're underway," probably account for the initial jump in pitch in the first TLI of launch. (This reference is to the astronaut's states of "probable relief" and "probable joy" which are defined situationally in Appendix B.) Such emotions ran high in Project Mercury.

A drop of 30 Hz is coincident with "Little bumpy" and the rapidly approaching region of increased vibration and acceleration. "Smoothing out," "Feels good," and "Through max. Q" are descriptive of the conditions accompanying the rise over TLI 8 and 9. There is a momentary drop of 40 Hz with "Sky looking very dark outside," but this is quickly reversed as evidence of a normal insertion accumulates; a high of 230 Hz is reached at BECO, a major event in a successful launch.

After TLI 14, the trend is downward toward more typical levels with a small, short-lived increment at Tower Jettison, another important sign of proper sequencing. The loss of this escape mode may account for the dip that follows immediately (TLI 18). A similar dip occurs in TLI 24 when the pilot notes, "My pitch checks at −7 at your −3"—a possible indication of trouble, to which the Cape responds, "Roger, seven." The news that "Cape is go" and the pilot's response, "Cape is go and I am go. Capsule is in good shape," mark the rise in TLI 27, which is followed by a dip while the last critical events of launch are awaited.

"SECO, posigrades fired okay," "Turnaround," and "View tremendous" send PHH to about 203 Hz. The trend line settles down a little over the next 5 min (TLI 2 to 6 of initial weightlessness) as capsule checks are executed, but rises sharply again (TLI 7) when he contacts Canary to report, "Control check complete . . . . Everything go . . . . Capsule in fine shape." Another slow decline,

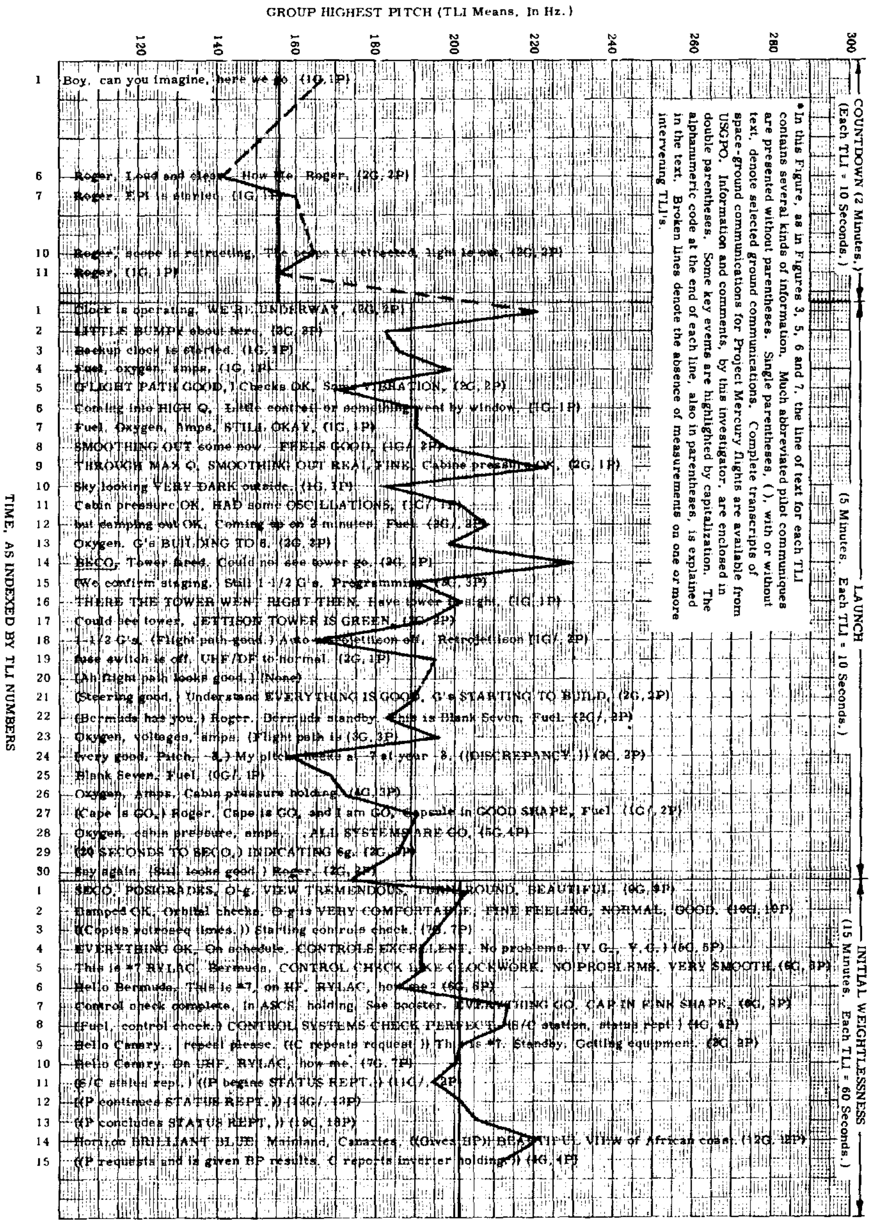


FIGURE 2.—Variations in Group Highest Pitch over a countdown, launch, and period of initial weightlessness.

here involving 25 PHH measurements, is spread over 4 min (TLI 8 to 11). Pitch builds again over the next 3 min as the pilot notes, "Horizon brilliant blue" and "Beautiful view of African coast."

"Probable relief" and "probable joy" (Appendix B) are the major elements of the pilot's state on attaining orbit. Their greater intensity here is due to several factors: the great significance of this achievement as a sign of mission success; initial weightlessness is "extremely comfortable" and "so pleasant it tends to become addictive," according to pilots; and the concerns ("probable apprehension") and physical stresses ("probable discomfort") of retrosequence and reentry are remote—hours of relative safety away.

Heart rate throughout these phases, although above average and somewhat responsive to the events of TLI's 1, 4, 8, and 14 of launch, shows general decline after the point of maximum physical effort (TLI 8) while pitch is still rising.

The differences between phase means—156 Hz for countdown (C), 189 for launch (L), and 202 for initial weightlessness (IW), indicated by the respective horizontal lines—are another salient feature. The corresponding standard deviations are 13 Hz for C, with N of 7; 21 for L, with N of 53; and 20 for IW, with N of 121. The t ratios, computed without pooling variances, were 6.8 for (L–C), 8.4 for (IW–C), and 3.8 for (IW–L), compared with 2.5, 2.4, and 2.0 for the associated Cochran-Cox (ref. 24) t' values computed at the 5-percent level.

To the extent that the trends already noted exceed the "noise," the t' test is to be doubted. If it is argued on the other hand that variations within phases are random and that the situational differences among phases are the prime determiners of state, the principal reason for doubting the t' test is removed. The former seems closer to the correct position.

Pitch perturbation during a countdown—An unusually high standard deviation of 37 Hz for one pilot during the 2-min countdown phase was held suspect at first, and the four PHH's contributing to it were checked; they were 179, 162, 63, and 151 Hz. The last two were from a single "sentence" with a hesitation pause of 0.19 sec between the subject and the verb. His voice apparently broke as he started the sentence, and a split-second pause was sufficient for regaining of full voice control—another illustration of fleeting emotional expression in superbly integrated and self-disciplined humans.

Pitch variations during retrosequence and reentry—These are shown (fig. 3) for a pilot:

(1) who, with less than 1 min to retrofire, had reason to believe that his clock was off by several seconds, signalling a potentially serious recovery problem

(2) whose intense efforts to establish ground communications just prior to this phase were unsuccessful

This is a clear instance of "probable apprehension" (Appendix B) for exceedingly good reasons.

The fact that PHH peaked after establishment of contact, while heart-rate standard deviation and speech rate peaked two TLI's earlier, may be further evidence of a relation between high PHH and "probable relief." Onset of the 30-sec retrowarning light early in TLI 4 probably accounts for the slight dip in PHH, tempered by the good news that "Retro attitude is green."

The three retrorockets fire during the first half of TLI 5 with some evidence of relief in the text of communiques for the rest of this TLI and the following one—perhaps the source of the successive increments in average PHH over the two periods. The PHH then declines over the next 4 min as "Yaw keeps banging in and out" ending with a sharp dip (TLI 9) where he decides, "I'll just control it manually."

Then comes the moment (TLI 10) when he is instructed to "Leave retropackage on through reentry." When he asks, "What is the reason for this?" he is told, "Not at this time; this is the judgment of Cape Flight." To this he replies, "Roger. Say again your instructions, please. Over." This presentation of any intense anxiety-producing stimulus [designated S^{-r} in the definition of "probable apprehension" (Appendix B)] in combination with a textbook example of an anger-producing situation (see "probable anger" in Appendix B) which is impossible to duplicate in a laboratory, is accompanied by a brief rise to 179 Hz (TLI 10) followed by a drop to the vicinity of 160 Hz where it remains (with low variation) for about 4 min (TLI 11 to 14), rising slightly (TLI 15) with, "I think the pack just let go .... A real fireball outside."*

Pitch rises and then falls as he tries to communicate through the "blackout" [(TLI 1 to 19); reentry]. When he does get through (TLI 20) and is asked, "How are you doing?" his answer, "Oh, pretty good," is a mild reflection of the terrifying uncertainties of a few moments before. Pitch is low in this region.

A rise begins with "Through peak-g" and con-

*The drop in PHH over the latter half of this phase was checked by the nonparametric Mann-Whitney U-test, the hypothesis being that TLI's 8 to 15 were lower in mean PHH than TLI's 1 to 7. (TLI 8 was assigned by chance.) The 0.001 significance must be treated cautiously since sample size varies over TLI's, so that this test is more appropriate to sequences of individual PHH's.

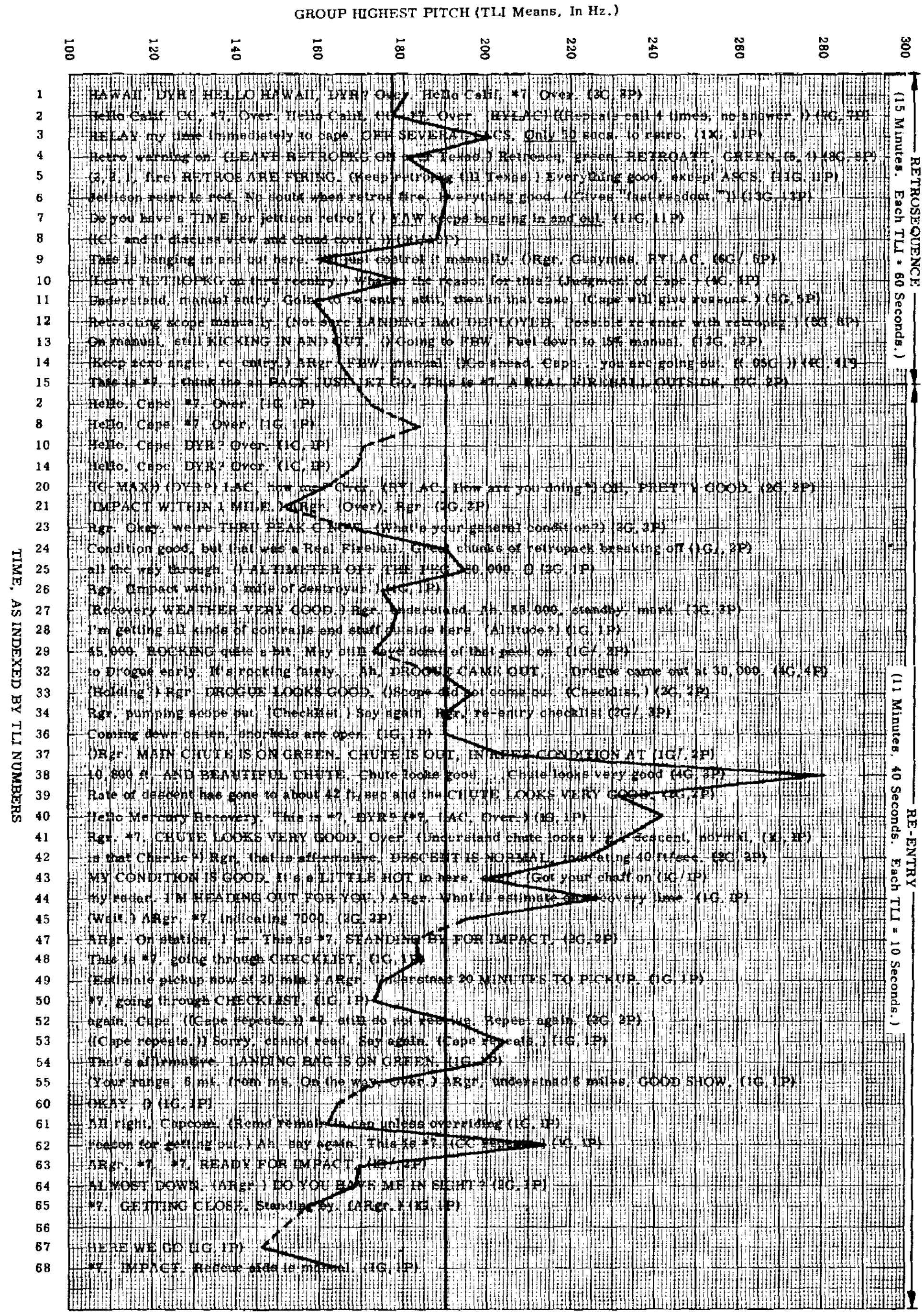


FIGURE 3.—Variations in Group Highest Pitch over a retrosequence and reentry.

tinues through TL25 when he reports "Altimeter off the peg 80 000"—the start of more-familiar altitudes. There is a drop over the next 50 sec (TLI 26–30), coincident with "contrails and stuff," the "rocking," and "I can't damp it either."

The intense joy and relief occasioned by the appearance of the main chute in these circumstances is eloquently expressed in the repeated affirmation, "chute looks good," and "beautiful chute." A skillful ground communicator recognized the intensity of feeling with, "Roger. Understand the chute very good . . . ." Meanwhile PHH skyrocketed to slightly over 280 Hz, a reading supported by additional oscillographic and spectrographic evidence. The decline that follows is slow.

Pitch variations during retrofire in another flight—Figure 4 is a pitch plot of a different sort, based on the oscillograph roll spanning the firing of retrorockets in another flight. Successive communiques and their subordinate Pause Groups are identified along the abscissa by evenly spaced index numbers. Each point on the plotted line is the PHH of the single Group for which the verbatim text is given. Relations between pitch changes and events are portrayed more precisely here despite some distortion in the time dimension. The TLI's are marked by vertical lines that intersect the text at the two places where they split Groups—in Groups 09-3 and 14-6. The five TLI means, which would have appeared as five successive points if plotted on a TLI basis, are here represented by horizontal lines.

The most striking point is the behavior of PHH during retrofire. The firing of retrorocket-1 is reported with a PHH of 179 Hz; retrorocket-2, with 201 Hz; and retrorocket-3, with 214 Hz. This is not a "cheer-leader pattern" with progressively higher pitches on "one," "two," and "three," but an apparently spontaneous expression of "relief" and "joy" at having "got three"—with the 214 Hz peak over the "got" on the oscillograph.

Also of interest is the alternation that is evident between a tendency toward low PHH in situations involving a countdown or wait for a critical event (recall the low pitch for the countdown phase of fig. 2) and a tendency toward higher pitch in reporting of the outcome: PHH drops from 217 Hz to about 120 Hz in anticipation of the 30-sec light, and rises to 147 Hz when it is reported; it drops as far as 113 Hz as he counts for the 20-sec light and tone, and rises to 180 Hz when he reports them. The countdown to squib arm (at 5) and to sequence (at zero) is more complex: there is a small drop as he notes that "Tone is out" and replies "That's correct" to CC's instruction to "Arm the squib at 5." However, his silence during the count provides no pitch measures until he announces the culminating events—"Squib arm" and "I have sequence . . . ."—which evidence a rise. The PHH exhibits another decline prior to retrofire and rises progressively as the rockets are announced. The reversal of the upward trend, in connection with the comment between the second and third rockets, also qualifies as an instance. The PHH drops prior to "Retrojettison armed" and rises when it is announced. A favorable fuel report momentarily reverses the downward trend toward "Retrojettison" which terminates in a sharp rise when the event is noted. A similar fall and rise occurs in the neighborhood of "Scope retracts." It is regrettable that time did not permit following this pilot through the reentry phase with this detailed analysis, or establishing the mean and range of his pitch excursions in other circumstances.

Figure 4 is methodologically important; it demonstrates that single values of pitch, each representing a segment (Pause Group) of a communique marked off by the 170-msec criterion silence and without adjustment for speech rate, can track key events—with some "noise" granted. This is short of surprising only because of the more-striking correlations of the figure 3 reentry which, after all, had an average of only 1.5 PHH's per TLI. The TLI means (horizontal lines) do reflect the low pitches during the countdowns preparatory to retrofire, the increase during retrofire, and the decline that follows. However, if the boundary of TLI 4 had occurred between 06-2 and 06-3, say, instead of between 05-1 and 05-2, the picture would have been blurred by a plot of TLI means. The standard deviation (not shown) being highest in TLI 5 alerts one to the fact that something has happened, but only observation of the details of figure 4 leads to any real insight.

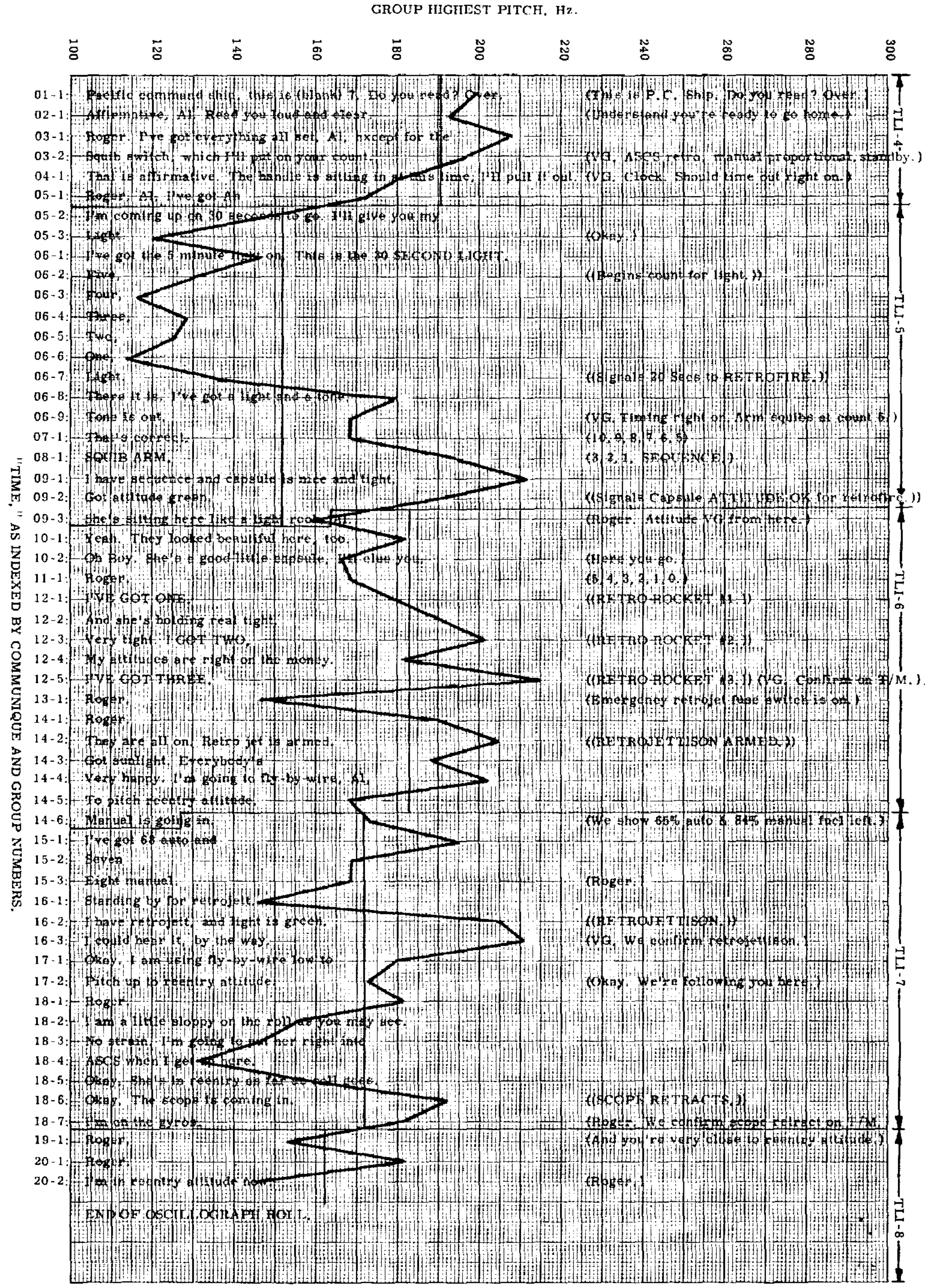


FIGURE 4.—Variations in Group Highest Pitch during a retrofire.

Some Variables Affecting PHH during High Arousal

What are the controlling variables causing pitch to be sometimes high and sometimes low when situational and other evidence indicates high arousal?

Low pitch (*PHH*) *during high arousal*—Several hypotheses are considered. First, individuals attempt to bring pitch down to levels clearly associated with calm, confident, competent performance in situations where important reinforcers are contingent upon demonstration to a listener that all variables are under control. Pitch inflections, intensity, pausal phenomena, and other discernible speech properties may be brought under similar control.

The plausibility of this hypothesis is suggested by (1) the existence of speech-aural and speech-kinesthetic feedback loops that provide the information essential to development of self-editing and self-controlling response; (2) the existence of reinforcement contingencies that foster the development of self-control more with respect to the expression of unpleasant or negative emotions than of pleasant or positive ones; and (3) the occurrence of instances of obvious control: for example, the quick resumption of normal speech by the pilot whose voice broke momentarily during a countdown, and the low pitch level maintained during the several minutes (TLI 11 to 15 of retrosequence) spanning the landing-bag discussion.

Second, an Estes-Skinner (ref. 25) conditioned-suppression paradigm is an element of situations involving countdowns or waits for critical events. The probability of a highly noxious or troublesome outcome, on termination of the count or time interval, is non-zero, and many of the important variables determining the outcome are beyond the control of the pilot. Experimentation is needed to determine whether it is proper to speak of the lower pitch observed in these circumstances as "pitch suppression," thereby linking it with the well-established phenomenon of "response-rate suppression." The fact that speech output is diminished during countdowns is compatible with this hypothesis.

Third, the on-going activities during countdowns and similar situations may explain the brief, intermittent utterances and the overall reduction in speech volume without appeal to conditioned suppression. During periods when fine-grain or numerous corrections must be introduced, or when an optimum state of readiness is needed for perception of multiple or near-threshold signals and making of appropriate responses, speech output is probably restricted to those minimal utterances (e.g., "Light is on" and "Tone is out") that were established in simulators as integral parts of the *S-R* chains. Routine acknowledgments ("Roger," "That's correct") and well-established verbal sequences ("5, 4, 3, 2, 1, 0") also may occur in these circumstances. The pilot's control over these multiple performances is probably maintained by discriminative responses to deviation signals as he scans his own outputs. When the system is relatively stable, the pilot can switch to generation of more-spontaneous messages such as "Oh boy. She's a good little capsule, I'll clue you."

However, these considerations do not identify any process by which pitch is lowered. On the contrary, increased glottal tension would be expected from the spread of muscular tensions associated with control of capsule attitudes and maintenance of a high degree of readiness to respond discriminatively, and from the task stress induced by concurrent critical activities.

Fourth, in a situation where the set of possible outcomes produces anxiety, the introduction of required task performances may reduce emotion by distraction (i.e., by substituting the emotional responses elicited by the inserted tasks) if one assumes that any strongly aversive possibilities present are not highly probable.

Finally, glottal relaxation may be an integral part of the "orienting reflex" elicited by the sequence of discriminative stimuli culminating in the critical outcome and mediated perhaps by such physiological correlates of this reflex as reduced heart rate and blood pressure.

Elevated pitch (*PHH*) *during high arousal*—Situations conforming to the definitions of "probable relief" and "probable joy" (Appendix B) yielded the highest pitches observed, for example, the attainment of orbit, the appearance of the main chute, and the favorable terminations of countdowns and waits. The greater freedom of expression accorded pleasant emotions is another factor.

High pitch may also be expected with unpleasant arousal under some conditions despite feedback loops. There is one such situation when the individual is attempting intensively and constructively to cope with a hazardous situation under time stress, and when the hazards are (or are thought to be) recognized by others; Berry's public comment on the high pitch of two Gemini astronauts while freeing their capsule from the perilously oscillating rendezvous booster is an illustration. Control of voice to communicate calm self-control would have been incongruous, particularly since the information communicated was enhanced by the pitch elevation. Another such situation is when arousal reaches the panic level, i.e., when the individual's behavior comes to resemble the energetic, problem-irrelevant activities of some animals in a conditioned-suppression procedure; and a third prevails when the immediate social environment is not so highly constraining as this one.

Finally elevated pitch is likely in communications under low signal-to-noise ratios, as discussed later.

The DURLL Ratio

Computation and significance—The end of a phrase is signalled by increased duration of the last syllable and by fine-grain variations in its pitch and intensity contours. The hypothesis that the acoustical correlates of linguistic terminals vary with changes in state of the speaker is most readily tested with syllable-duration data. To the extent that Pause Groups and phrase boundaries are coincident, the final syllables of Groups tend to be longer than the same phonemic combinations occurring elsewhere. A strong association between the two is expected in astronauts' speech since, as noted earlier, most of it is rehearsed, descriptive, or concerned with matters extensively discussed before flight.

DURLL, as the DURation of the group Last syLlable was called for FORTRAN programming, is the time between the onset of the last syllable of the Group and the end of the Group. DURL is the average duration of the remaining syllables, computed over a Group, a TLI, or a flight phase. The DURLL Ratio = DURLL/DURL. This normalization should be made on a Pause Group basis; instead, as one of the compromises of the "economy size" computer program, TLI means were used.

The DURLL Ratio during countdown, launch, and initial weightlessness—The three horizontal lines in figure 5, at values of 0.99, 1.40, and 1.39 for phases C, L, and IW, respectively, are the mean ratios computed from the corresponding light-phase means for DURLL and DURL. The lines are obviously not centered on the trends, since the ratio of the means is not equal to the mean of the ratios.

The dramatic increase from lift-off through BECO and on to a peak of 2.26 at Tower Jettison may well reflect the concern in Project Mercury with these early stages of the launch phase, and the joy and relief on their successful completion. This increase is not a function of acceleration since the line continues to rise as acceleration falls after BECO, and declines during much of the subsequent buildup of gravity (TLI 18 to 26). The downward trend is reversed with "Cape is go . . . . All systems are go" and the approach of SECO.

A peak of about 1.58 (TLI 2 of IW) follows SECO, turnaround, and the attainment of weightlessness. The value declines over the next 7 min (TLI 3 to 9) during capsule checks, and then rises sharply as he begins his very favorable report to Canary.

The DURLL Ratio during retrosequence and reentry—The downward trend over the 11+ min from TLI 6 of retrosequence to TLI 14 of reentry is clear in figure 6. The values over the 4+ min from the end of the communications blackout to beyond the opening of the chutes (TLI 14 to 41) are puzzling. They are not only low but show their peaks one or two TLI's after the corresponding ones for pitch: TLI 25 for PHH versus TLI 27 for the DURLL Ratio; and similarly 33 versus 34, 38 versus 39, 44 versus 47, and 53 versus 54. The dip at the end of phase IW (fig. 5) while PHH is peaking (fig. 3) is also noted.

The most plausible explanation is that pitch prominence is associated with the lengthening of syllables earlier than the last in sequences like "Brilliant blue," "A real fireball," "Tremendous view," and "Beautiful chute," resulting in temporary depression of the DURLL Ratios in these TLI's. If so, substitution of the Group Maximum Syllable Duration for DURLL, as discussed later,

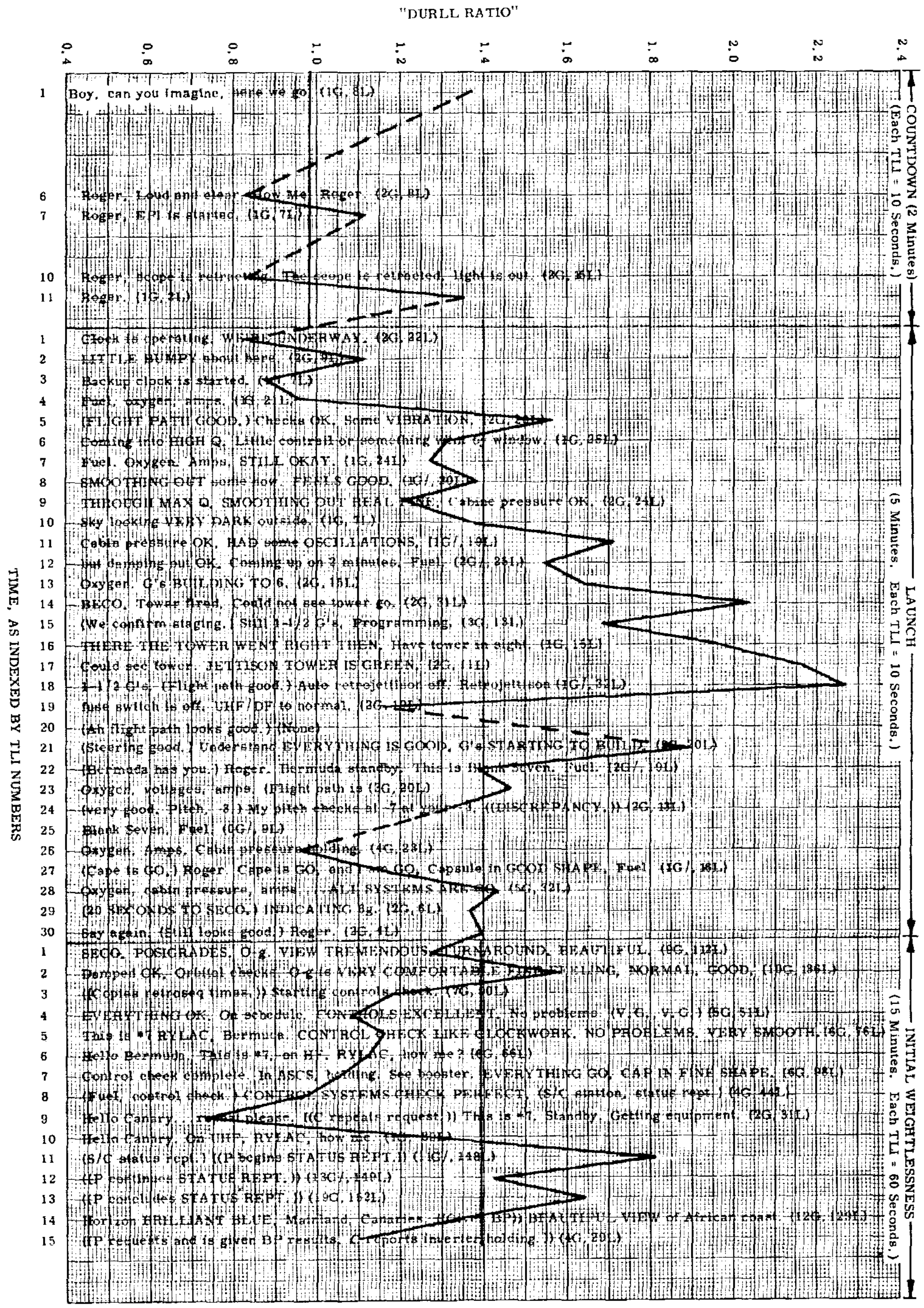


FIGURE 5.—Variations in the DURLL Ratio over a countdown, launch, and period of initial weightlessness.

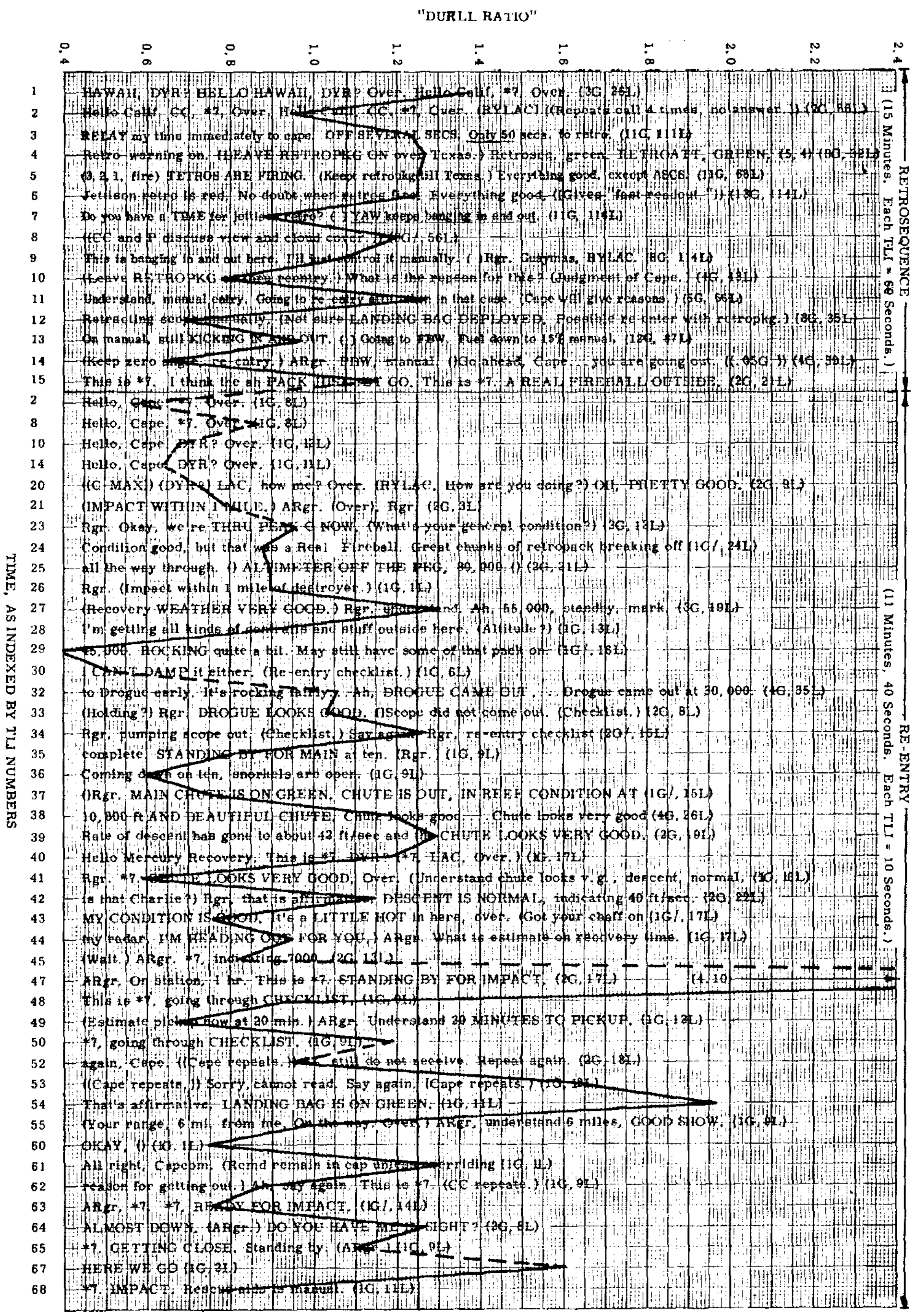


FIGURE 6.—Variations in the DURLL Ratio over a retrosequence and reentry.

would correct for this and should be tried. Segmentations not coincident with phrase boundaries, due to hesitation pauses or the use of a fixed criterion silence regardless of speech rate, also could produce low values. Finally it is possible that the low values reflect the state of the astronaut after what must have been an emotionally exhausting flight.

Concern about the false landing-bag-deployment signal, telemetered from orbit, is a likely explanation of the extraordinary value of 4.10 (TLI 47) when he says "This is [blank] seven, standing by for impact." This is followed by another peak (1.96, TLI 54) when he confirms, "Landing bag is on green"—an important S^{+r} (Appendices) even if it did not dispel all uncertainties.

The DURML Ratio as an alternative—The possibility of substituting the DURation of the Maximum syLlable (DURML) extent for the DURLL of each Group, mentioned above, was discussed in the project report. The measure has considerable promise. It is less dependent on the assumptions of coincidence between Pause Group and phrase or sentence boundaries; in fact the DURML measure would be improved by an increase in the criterion silence to perhaps 250 msec. The same is true for Group Highest Pitch but probably not for the DURLL Ratio. It reflects the expressive lengthening of syllables other than the last, as in "... *beau*tiful chute" and other examples cited above. It is equivalent to the DURLL Ratio when the Group's last syllable is in fact the longest.

On the other hand, the DURML Ratio mixes two different linguistic processes and two different definitions of the syllable in the numerator. The respective points can be identified in computer plots, and it is possible that a level correction may be derived to remove discontinuities between the two. DURML is typically more dependent on the precision of syllable marking, being bounded by syllable onsets. DURLL requires only one syllable-onset, but is subject to error in detection of the end of the Group in noisy speech.

Syllable Peak Amplitude

Syllable Peak Amplitude in volts rms (ARMS; fig. 7) seems more closely related in time to indications of apprehensive arousal than is pitch (fig. 3). ARMS starts higher (relative to the retrosequence mean) than PHH and continues to climb until retrofire is over; then it drops markedly. Pitch, on the other hand, begins to decline as soon as contact is reestablished and assurances are received that retrofire time will be signalled properly. The initial association of elevated pitch with high ARMS may be another illustration of reduced voice control while an obviously urgent situation, as seen by the pilot, is coped with, or of the use of these parameters to communicate urgency. These interpretations are reinforced by the peaking of heart-rate variability in TLI 2. Or the higher subglottal pressures, generating the loud speech by which the pilot attempted to reestablish communications, may have induced glottal tensions, thereby increasing voice frequency. This process may well contribute to the high ARMS and PHH in the vicinity of TLI 62 where channel conditions occasion "Say again." However, pitch does not rise when the pilot shouts (TLI 14) to get through the communications blackout of rèentry, and it falls as ARMS rises prior to retrofire (as already noted), suggesting that this process is at most a small part of the observed variations.

ARMS rises more dramatically than does pitch in response to the order to leave the retropackage "on" and does not diminish, as pitch does, when the request for an explanation is denied. The hypothesis relating low pitch during this period to the pilot's control of intense emotions is recalled.

There was closer agreement in general contours, when the plot of TLI heart-rate means was superimposed on the ARMS plot for retrosequence, than with PHH. Heart rate starts moderately high, rises a little to retrofire in TLI 5, and declines about 15 beats per minute (bpm) by TLI 8; it starts to rise in TLI 9 with "This [automatic yaw control] is banging in and out here," and continues to a peak slightly above 100 bpm (with rates within the TLI reaching 120) when the explanation is finally given in TLI 12. It then declines somewhat over the next two TLI's and rises again in TLI 15.

The lag between heart rate and ARMS is less than that between heart rate and PHH when these variables peak in response to events culminating in appearance of the main chute. The ARMS

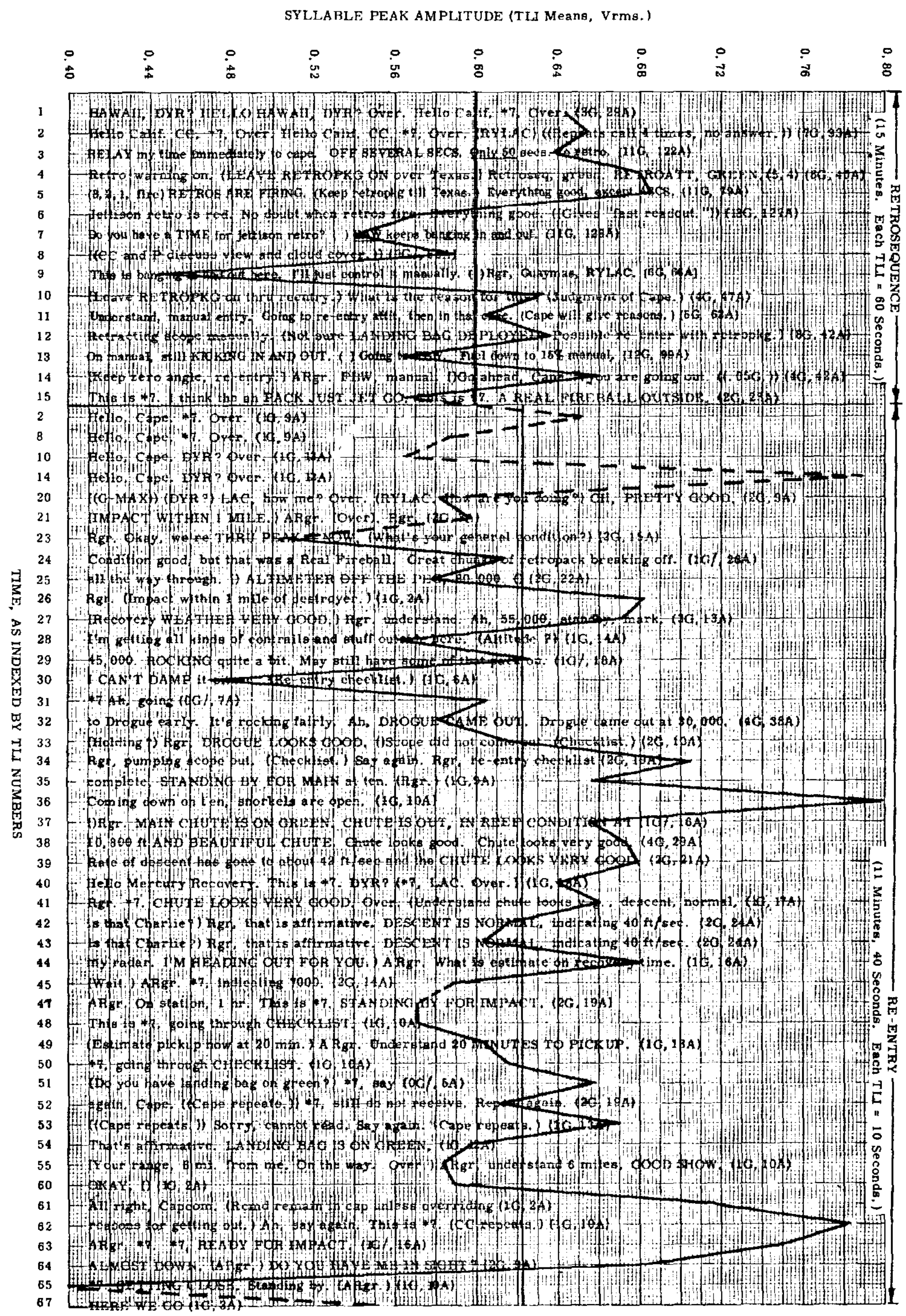


FIGURE 7.—Variations in syllable amplitude peaks over a retrosequence and reentry.

shows its major peak in TLI 36, declines sharply in TLI 37, and evidences only a small increment when PHH attains its maximum with a lag of 20 sec in TLI 38. Heart rate peaks (138 bpm) in TLI 34, drops slightly by TLI 35, is about 120 bpm when ARMS and PHH show their respective peaks, and continues its decline until TLI 45 where it rises slightly, with ARMS and PHH trailing it all the way in that order. This lag in PHH is compatible with previously noted associations between the "relief" and "joy" combination and pitch peaks.

These observations are merely suggestive since time plots are intended mainly for demonstration of the responsiveness of parameters to situational changes. Relations among variables and the assignment of "probable state" labels are more properly studied when all data are assembled on a Pause Group basis for analysis, as discussed in the Appendixes.

The tremendous drop (0.68 to 0.38 V rms) between TLI's 64 and 65 is without parallel, perhaps involving capsule power, bracing for impact, or other physical cause. The value recovers to 0.55 V rms in TLI 67 but dives again to 0.38 V rms in TLI 68 (not plotted) when he says, ". . . seven. Impact. Rescue is manual."

Syllable Rate

This measure was computed as the reciprocal of each Syllable Duration before averaging over a TLI. This method of computation assigns higher weights to faster syllables since the reciprocal function decreases rapidly as Syllable Duration increases from below 0.1 to about 0.3 sec. This differential weighting probably explains the apparently greater validity of Syllable Rate (RL), as compared with Syllable Duration, both graphically and in the statistical checks employed. In this respect, RL is akin to the widely used beat-by-beat index for heart rate. The question of whether the reciprocal transformation is appropriate to the respective mean-variance relations was not investigated.

The duration of the last syllable in each Pause Group (DURLL) was excluded from RL, being under the control of unique linguistic constraints. The reasonableness of this choice is illustrated by the first Group uttered in TLI 1 of the retro-sequence, where the pilot was desperately trying to establish communications before retrofire: "Hawaii, did you receive? Over." The strong motivations of the moment were reflected both by the extended durations (0.22 and 0.32 sec, respectively) of the two syllables in "receive" and by the rapid articulation of the other syllables at a mean of 0.14 sec. Inclusion of DURLL (or DURML) in rate measures would tend to obscure these subtleties.

It is interesting that the mean TLI Syllable Rate for the pilot most extensively studied ranged from 4.0 to 8.0 syllables per second on the basis of 24 and 23 syllables in the respective TLI's. One lower mean of 2.8 occurred, based on a single syllable. Differences between these results and typical published figures are due to the elimination of DURLL, the taking of reciprocals, the elimination of silences greater than 170 msec, and the relatively small number of syllables in each computation.

First-Order and Second-Order Difference Measures

Hypothesis and specific measures—On the assumption that the sequential properties of the number series, representing a given measurement on successive syllables, might vary with the state of the speaker, algebraic first-order and second-order serial differences were computed for Syllable Duration (DL), RL, and Syllable Peak Amplitude (for which A is an abbreviation of ARMS). The standard deviation was selected as the most promising measure. With SD symbolizing this statistic and F1 and F2 used for finite differences of first-order and second order, the variable names of interest are SD(F1RL) and SD(F2RL), SD(F1DL) and SD(F2DL), and SD(F1A) and SD(F2A). In continuation of this symbolism the respective standard deviations, computed without regard to sequential dependencies, are SD(RL), SD(DL), and SD(A). Omission of the Group Last Syllable from indices involving RL and DL probably diminished the validity of these measures a little, but reduced redundancy with the DURLL Ratio to which they are somewhat related.

Some Properties—Inspection of the time plots (not presented here) indicated that (1) the SD(F1) and SD(F2) measures tended to covary for each parameter; (2) the difference measures for RL and A showed some good correlations with flight events; (3) the lower apparent validity of SD(F1DL) and SD(F2DL) is further evidence that the occurrence of rapid syllables within a Pause Group is a more sensitive index of certain psychophysiological stresses than average syllable rate or duration, as ordinarily computed; and, just as RL weights these shorter syllable durations more heavily than does DL, by virtue of the reciprocal transformation, so SD(F1RL) weights successive differences involving shorter durations more heavily than does SD(F1DL); and (4) the TLI plots for SD(F1RL) and SD(F2RL) showed occasional, sudden, extreme values or "blow ups."

To clarify the properties of these difference measures, a simple model was constructed consisting of strings of U's and A's (Unaccented and Accented) related by a Contrast ratio (C) greater than 1 so that A's are uniformly higher in value—longer, more intense, or higher in pitch. Expressions for SD(F1) and SD(F2)/SD(F1) were found to converge rapidly on $\sqrt{3}$ as a lower bound as the number of "syllables" increased.

Considering the simplicity of the model, it was a great surprise to find the actual values for all variables and phases of flight (ranging from 5 to 15 min) in close agreement: the mean for 17 phases (exclusive of prelaunch countdowns) was 1.73, with 16 values ranging from 1.68 to 1.77 and one value for a launch at 1.88 (for DL). The averages for the three variables were 1.715 for RL, 1.738 for DL, and 1.740 for A (ARMS). For two separate prelaunch countdowns the values were 1.44 and 1.46 for DL and 1.51 and 1.57 for RL, a result that merits further study. The corresponding values for ARMS were 1.80 and 1.69.

Furthermore, when the computed values for SD(F1RL) were multiplied by $\sqrt{3}$ as a scale factor and plotted, the correlation with SD(F2RL) was nearly perfect except for several excursions of SD(F2RL) where it seemed to contain additional information. Attention was therefore focused on the expression for SD(F1RL), which is simpler. The constraints were redefined, and the expression was modified to cover either phrase-like strings or indefinitely long strings that approximate a little more closely those of spoken language.*

Several examples of "blow ups" in SD(F1RL) and SD(F2RL) were also examined, one being the now-familiar first TLI of the retrosequence which includes the two consecutive Pause Groups whose syllables and durations are as follows: "HA- (0.06 sec) WA- (0.17) II (0.14), DID (0.13) YOU (0.13) RE- (0.23) CEIVE (0.22)? O- (0.15) VER." (0.19 = DURLL, i.e., last syllable of Pause Group #1); and "(H)E- (0.10) LLO (0.19) HA- (0.13) WA- (0.17) II (0.14), DID (0.15) YOU (0.13) RE- (0.22) CEIVE?" (0.32 = DURLL).

As suspected, it is the taking of reciprocals of occasional, very short syllables (e.g., 0.06 sec) and the squaring of successive differences that cause the trouble, affecting SD(F2) more than SD(F1). The ratio SD(F1RL)/SD(RL) virtually eliminates "blow ups" and provides a conceptually better measure of sequential dependencies by normalizing against the level of nonsequentially dependent variability. Time plots of this ratio showed responsiveness to situational changes, but this is an area requiring more research.

Other Speech Parameters

Pause Group Duration (DURG) and Syllables per Group (ENLG) showed some promise in the TLI plots; ENLG appeared more responsive, but DURG made a stronger showing in the statistical comparisons made. The Coefficient of Variation suggested that ENLG lost out in flight-phase comparisons where it was more variable—most responsive. The values of this coefficient were themselves interesting; for example, DURG was lowest (48) during a countdown and highest (74) during a retrosequence. To the extent that inhalations occur during criterion pauses, ENLG approximates Goldman-Eisler's (refs. 11 to 13) Syllable Expulsion Rate (ER) computed as syllables per expiration. Under conditions of "reasonably constant sound pressure level" she interprets the reciprocal, 1/ER, as the proportion of returned air current used to produce each syllable. Topical analyses and "correlation" with

*The luxury of pursuing such questions (without regard to contractual scope, economics, or other practical values) derived from the fact that most of the statistical and graphical analysis of data for this project was completed by me in my own time out of personal interest. The computer provided the contractually required phase and TLI means and standard deviations, together with sums of squares and a convenient ordering to facilitate manual analysis.

some psychiatric ratings suggested that "high values belong to content implying free flowing or outgoing effect" and "low indices . . . belong to topics of restricted emotionality, which implies tension states or intellectualized speech."

Duration of pausal silences (inter-Group silences greater than the 170-msec criterion pause) was poorly distributed over TLI's and showed no promise of validity.

RECOMMENDATIONS

A Semiautomatic Monitoring System

General description—Detection of changes in speech parameters by human monitors would be enhanced by a real-time, CRT display of functions such as $S(t)$, $A(t)$, and $P(t)$ (fig. 1).

On detecting or suspecting a change in state, the monitor would initiate analysis of selected speech samples. This would be done at a console equipped for oscillographic readout, for computer access via an x-y oscillograph digitizer capable of coordinated keyboard entries in an efficient code, and for playback over a system such as the one used in this research, which enables repetitive audio and CRT display of short speech selections without the inconvenience of tape loops or the long recycle times of Tape Search Units.

Two levels of scoring are envisaged: Measures in level-1 would be selected by further research for both validity and speed. The raw inputs might include: (1) Group "on" and Group "off" times encoded as distances along x from a reference point; (2) the onset of the DURLL or, if a different DURML were readily apparent, the two onsets bounding it; (3) PHH and its calibrated base line, P(60 Hz); and (4) a Group Syllable Count derived by tapping of the perceived syllables into a digital counter, perhaps with checking of the oscillograph to avoid the error of perception of elided syllables. A less precise count, probably adequate for level-1 scoring, can be made directly on the oscillograph without auditing of the communique.

With these simple inputs, time plots (preferably on a Pause Group basis as in fig. 4) can be generated by a small computer for the following variables: (1) Group Duration (DURG); (2) a DURLL Ratio, modified to permit either DURLL or DURML in the numerator and normalized against an average syllable duration (DURL) derived from the Syllable Count and the Group Duration with the last (or longest) syllable excluded from both; (3) PHH; (4) RL; (5) Syllables per Group; or (6) a validated selection of these. This is a vast simplification of the procedures used in this research, particularly with respect to syllable marking.

Level-2 scoring would be applied whenever a more detailed analysis seemed warranted by level-1 results, medical data, or other on-line information: (1) Syllable Peak Amplitude (ARMS) is a promising example, but requires a normalization to correct for level changes during propagation and elsewhere; this could be accomplished by digitizing of each peak as we have done; then, instead of computing the average, the computer would select the Group's highest peak and normalize it against the average of the remaining ones, forming an AHH Ratio analogous to the DURLL and DURML ratios and (except for normalization) to PHH. (2) A computer operation on the same amplitude data could yield the difference measure SD(F1A)/SD(A). Finally, if the marking of syllables could be accomplished or facilitated automatically or if the considerable labor of manual marking could be justified, (3) the DURLL or DURML Ratio and (4) the RL measures could be improved; and (5) the difference measure SD(F1RL)/SD(RL) on LR could be added.

Requisite research—Assembly of a laboratory version of the proposed operational console, with computer access, and its use for developing procedures and for conducting research is recommended. Crews of Apollo, MOL's, or high-performance aircraft could serve as subjects in simulators or in flight. Some development and improvement of speech-processing equipment would be necessary for the recommended traces and for promising new ones.

Criterion pauses should be studied as a function of speech rate, with the results applied to segment communiques into Pause Groups—for example, by employing a very short value (such as 150 msec) and programming the computer to apply and regroup entries by the longer criteria derived as a function of speech rate.

Other questions merit research: Does segmentation by a rate-proportionate criterion pause coincide any better with listener-detected pauses? With breath pauses? (A respiration time function recorded on a channel parallel with speech would be desirable.) With linguistically appropriate boundaries? Does the distribution of articulatory pauses vary with emotional state when normalized for speech rate? How do these variations affect the optimum criterion pause which is essentially an upper limit on or at least a very high value within this distribution?

To what extent would a combination of criteria—for example, the requirement of an increased DURLL Ratio on the syllable preceding a criterion pause—improve segmentation? Improvement appears likely, since syllable prolongation is one of the linguistic "terminals" that signal the end of a phrase (ref. 26), even one interrupted by hesitations. With good-quality speech, the addition of terminal pitch and amplitude criteria would merit consideration.

Exploratory validations should be continued, utilizing time plots, trend tests, and other simple alternatives to Probable State Analysis. Finally, when procedures are stabilized and measures have been selected, a full-scale validation study should be executed, using a "streamlined" version of Probable State Analysis (Appendices A and B).

Feasibility of a Wholly Automatic System

The main hope for an automatic speaker-state monitoring system, applicable to space flight within the next 5 years or so, lies in development of a spacecraft channel vocoder (ref. 27), and computer programs for analysis of the digital transmissions to identify Pause Groups, Group Highest Pitch, and other parameters including some new measures not possible with current analog transmissions or on-board recordings. In fact, programs of this type—for example, for use of spectral information for improvement of syllable-onset detection during voiced segments, particularly in the presence of semivowels, voiced fricatives, and nasals, and for formant tracking—have been developed in my laboratory and elsewhere.

The speech synthesized at the receiving end of a channel-vocoder system has an artificial quality, seeming to lack much of the aurally discriminable state information ordinarily important to medical monitors. On the other hand, this system preserves most automatically analyzable speaker-state information in addition to making better use of available power for transmitting intelligible speech over great distances. The trade-offs will be particularly clear if speaker-state analysis proves its worth in further validations.

Hybrid (voiced-excited) vocoders digitize a base band of low-frequency voice energy as an analog function and employ channel-vocoder techniques for the rest of the transmitted spectrum. At the receiving end, the narrow base band is distorted to yield a wide, flat spectrum capable of energizing all channels of the synthesizer; this improves speech quality and preserves more aurally discriminable information on state, but complicates automatic speaker-state analysis. Vocoders have been produced by General Dynamics, for example, with selectable channels and voice-excited modes.

Supplementary Laboratory Research

General—The search for speech parameters and weighting functions that discriminate among probable states would also be facilitated by a simple laboratory set-up for testing and clarifying relations in data from simulators and actual flight.

A set of standard individual and group tasks must be developed for generation of speech under laboratory control. The several classes of verbal responses, differentiated on the basis of controlling stimuli (ref. 28), should be represented. Some should generate "constrained" speech—for example, propositional functions or formats in which the subject inserts the appropriate word from a prescribed set; others should generate relatively "unconstrained" or "free style" speech. The verbal output should be an integral part of the task; thus the experimenter's interest in it is concealed. The tasks should facilitate application of various reinforcement contingencies for manipulating the state of the speaker.

Generation of "unconstrained" speech—A signal-detection task is probably the most efficient way to generate quantities of comparable verbal responses of the "constrained" type for each experimental condition. Experimental error variance is reduced by the use of two stimulus conditions—Signal plus Noise (SN) and Noise Alone (N)—and the requirement of only two verbal responses. Presentations may be initiated by the subject or the experimenter.

The S/N ratio may be decreased either to increase the difficulty of the task or to deprive the subject of a dependable basis for his decisions and permit administration of a random or planned sequence of successes and failures. However, this procedure shifts response control from the stimulus to sequential or other adventitious variables (ref. 29). The suggested inclusion (ref. 30) of easier discriminations in near-threshold studies maintains discriminative responding by a periodic reinforcement.

The basic verbal report may be Signal/No Signal, Affirmative/Negative, Yes/No, etc. In addition the subject may be required to request each stimulus presentation, to assign a "confidence level" to each decision, or to keep track of and report presentation numbers and outcomes for the record. The possibilities can be expanded by combining these with any one of a variety of simple numerical, geometric, or verbal tasks, while establishing contingencies on only the detection task.

Reduction in the number of decisions, and consequently in the value of signal-detection statistics as a source of information about the state of the subject, is the price paid for this greater variety of verbalizations and the greater task complexity; this is directly compensated by data on a more diverse range of task performances.

Provision for pressure-key responding—A pressure key was installed as part of the signal-detection setup that was nearly completed before this project arrived (refs. 31 and 32). The intention was to determine whether changes in the durational and intensive properties of speech during state changes are accompanied by analogous changes in the responses of a different musculature not subject to specific social conditioning and styling.

This requires freeing pressure-key response from the control of subvocal speech which occurs, for example, in tapping-out the syllables of "affirmative" and "negative," or in counting to press a specified number of times for "yes" or "no." Dot-dash patterns encourage the dit-dah verbalization of the Morse code beginner. A possible solution is to press a variable number of times (say 3 to 5) until a light indicating "affirmative" goes "off," and an independently varied additional number of times for a negative-decision indicator. This would bring key-press responding under the control of the respective indicator lights and remove or drastically reduce subvocal control which is essential to the inferences of interest.

Manipulation of speaker's state in the laboratory—The concepts and stimulus operations of reinforcement theory provide the main basis. Contingent reinforcements (those determined by the correctness of the decisions) can be accommodated by signal-detection theory to the extent that they can be expressed in a "loss function." Noncontingent reinforcements used to manipulate state are outside the formal structure of decision theory as applied to the detection task (but not to the subject's decision to leave or go along with the situation). The task lends itself readily to either type of reinforcement.

"Probable apprehension" (Appendix B) can be simulated by conditioning of a small panel light as a sign (S^{-r}) of impending, unavoidable, aversive stimulation at the end of about 3 min in accordance with the Estes-Skinner "conditioned suppression" procedure (ref. 25). Termination of the warning light by either an aversive or a pleasant outcome, in some probability mix, is more analogous to real countdowns.

A type of "task stress" can be induced by a modification of the Sidman avoidance technique (refs. 33 and 34). In the standard procedure the subject receives one shock for failure to respond within *n* seconds of his last reponse (the "response-shock interval") and an additional shock every *m* seconds (the "shock-shock interval") until he responds again; the parameters *m* and *n* and shock intensity are important. As applied to the signal-detection task, shock-avoidance must be made contingent not on the mere occurrence but on the correctness of decisions during the response-shock interval—for example, *q* correct decisions every *t* seconds, with a new interval initiated when the criterion is met or at the end of *t* seconds. With certain values of *q*, *t*, and S/N, the subject can increase the probability of postponement of the shock more by crowding extra responses into the response-shock interval than by taking extra time with each decision because of the nature of the task. The procedure would be superimposed on a schedule providing payment for each correct response. With knowledge of the distribution of task-cycle times for regular reinforcement, the probability of drawing a sample of n responses, with an average time not exceeding a specified value, can be used to establish t. The speed incentive can be removed by a modification, that may be termed the "block correctness ratio method," in which the subject must produce correct responses in each block of n to avoid aversive stimulation at the end of the block.

Boredom may be occasioned by long sessions, easy decisions, a flat hourly rate with no payoffs or punishments, or long waits between series. It can be relieved at least temporarily by more-complex contingencies.

These were among the paradigms examined and adapted for assessment of the versatility of the signal-detection task for investigation of changes of state relevant to space flight as they affect "constrained" speech. The applicability of these contingencies to some tasks suitable for generating "constrained" speech has also been explored. In addition, several experiments aimed at clarification of the speech-parameter variations observed during countdowns and waits have been sketched. An informal report (ref. 35) is available from the writer on request.

Evidence of the effectiveness of these laboratory "situational" changes would be sought in measurable changes in task-performance data, introspective reports and checklists, psychophysiological measures, and other response information.

ACKNOWLEDGMENTS

The electronic equipment was engineered by L. C. Stewart and W. D. Larkin, and the calibrated oscillographs were painstakingly produced by S. B. Swackhamer. Drs. R. A. Houde and W. B. Newcomb provided occasional consultation in linguistics, as did Mrs. Marian MacEachron who also helped with the computer programming. Dr. Willis marked syllable onsets and performed the tedious oscillographic measurements and numeric encoding in the summer of 1966 when he was a graduate student in linguistics at the University of Rochester. The favorable responses of Drs. C. A. Berry and A. D. Catterson, Manned Spacecraft Center, to a proposal in 1964 resulted in sponsorship by Dr. J. F. Lindsey who saw the possibility of incorporating the proposed "probable state" concepts and acoustical techniques with his Time-Line-Analysis approach.

Page intentionally left blank

APPENDIX A

ASSESSMENT OF ASTRONAUTS' CHANGES OF STATE FOR SPEECH RESEARCH

PROBABLE-STATE ANALYSIS

General—Probable-state analysis is a method for organizing situational and response evidence sequentially, and for using it for generation of 10 separate time functions—one for each of the probable states selected—depicting variations in the state of the pilot.

The rationale was described in a preproposal document (ref. 36) and further developed during this study. The modifier "probable" was introduced to emphasize the incompleteness of evidence ordinarily available and the consequent need to avoid direct and unqualified assertions about the state of a pilot—except perhaps where the evidence is written plainly for all to see.

Spacing and time span of state and speech measurements—The speech-pressure wave form is a time function, so are its various transformations into voltages representing amplitude, pitch, formants, and other parameters. These functions exhibit marked discontinuities within periods of so-called continuous speech and go to zero between communiques.

Measurement operations on these transformed functions yield sets of numbers or measurement vectors, each representing the pilot's speech behavior over the time span of the communique, or a segment thereof (Pause Group), from which the measurements are made. Speech segments are spaced unevenly over time, and measurements from them are identically spaced. Probable-state vectors must have the same spacing and time span as have the speech measurements if relations between the two are to be explored—a type of asynchronous "sampling" that may be called "event paced."

Fortunately the text of space-ground communications is a very valuable source of information about situational changes and pilots' responses to them, and the interval between an event and its report is usually short. These factors facilitate setting of state-related data into approximate time correspondence with speech measurements for analysis.

The Pause Group as a unit of analysis for validation studies—As reported in the main text, physical measurements have been made from the Group as a subdivision of the communique and from the syllable as a subdivision of the Group. For bringing all measurements into proper time correspondence for a validation study, those from the syllable must be converted to a Group basis; separate state vector must be generated for each Group. It is the state of the pilot over the time span of each Group, not the Group or its text *per se*, that is assessed in the light of all available evidence except the physical properties of his speech! This sets the stage for the statistical analysis of results and is quite compatible with the Time-Line Analysis approach. For, if successful, state predictions made from the acoustical measurements from each Pause Group could be assigned to the TLI in which all or the greater number of syllables occurred, or distributed with appropriate weights.

The main problem is the number of judgments required of raters. Summing of all acoustical measures over communiques and generation of state assessments over their respective time spans ease this task somewhat but add new problems. Communiques vary in length from a "Roger" to status reports running many minutes in orbital flights. Even if these are subdivided topically, the range is too great from the standpoint of statistical analysis or TLI application.

However, for testing of the procedures and concepts of probable-state analysis, this expedient was adopted. Data on state were organized sequentially by communiques or their topical subdivisions with the proviso that further subdivisions could be effected whenever necessary to portray changes in state that were considered significant. A suborbital flight was chosen for the trial analysis.

Sequential organization of probable-state information—An adaptation of the techniques of Barker (refs. 1 to 3), for recording and portraying the "stream of behavior" as a first step in analysis of it, was found most suitable for our purposes.

Figure 8, one of 14 worksheets required to represent a suborbital flight, illustrates the adapted procedures and format. A verbatim transcript of pilot-ground communications occupies the central portion. The pilot's communiques (*Q*'s) are in solid capitals. Left and right portions of the original 11×17-in. data sheets have been deleted since some of the information—introspections, retrospections, task-performance details, situational factors, and medical data—cannot be released without approvals.

The first of three concepts requiring discussion is Barker's "episode," symbolized here by *E*. "Although continuous, the stream of behavior occurs in perceptible units. One of these units is the behavior episode. Episodes are the 'things' that people normally see themselves doing [Eating

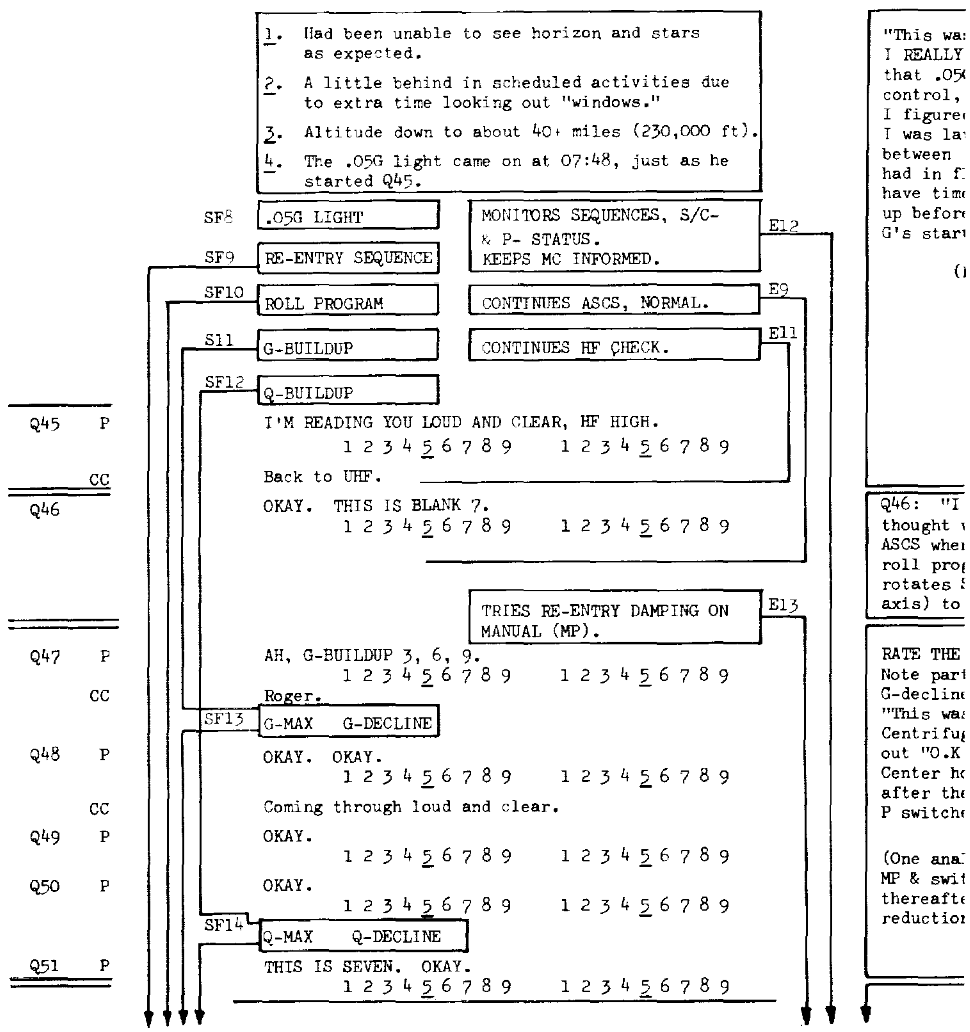


FIGURE 8.—Excerpt from a "stream of behavior" representation of a suborbital flight.

an apple and playing cricket are given as examples.] . . . the action of an episode is directed toward a single goal."

Barker and his co-workers prefer a format in which the central portion contains short paragraphs by on-the-spot observers, describing in everyday language the behavior of a single individual, usually in a social setting. An exception is the study by Soskin and John (ref. 37) who present the recorded conversations of a man and wife encountering river traffic in a rowboat; episodes are occasionally marked, but incidentally to their main objectives.

Episodes are indexed in order of occurrence, with new ones added as innermost brackets. For convenience, monitored sequences, reporting the status of spacecraft and pilot and maintaining communications, are grouped in a single block with the bracket extending the length of the phase.

Episodes correspond to a chain or series of responses under the control of a single but possibly complex set of reinforcement contingencies, or to a response sequence under the control of a "situation-and-goal set" in Woodworth's (ref. 38) phenomenologically oriented terminology.

Blocks and brackets on the left side of figure 8 denote "situational factors" (SF). Schoggen (ref. 39) calls these "environmental force units" (EFU's) (reflecting the Gestalt-Lewin background of this work) and defines them to include the physical and social process or events that "spur, guide or restrain behavior." In Mercury flights, SF's were mostly physical—lift-off, build up of gravity, BECO, etc.—with social interactions confined to communications channels.

The mere occurrence of an environmental change is not sufficient. A false landing-bag-deployment signal, for example, lacks the full credentials of an SF (or EFU) until the pilot evidences some "awareness" of it by his responses. Barker, in the Gestalt tradition, requires evidence that the change has "penetrated" the individual's "psychological world"; this parallels the behaviorist view that a change is a "stimulus" only if followed by a detectable response or an alternation in response parameters (ref. 40). The small subset of stimuli, singled out for charting, is determined by the framework and objectives of the analysis. In the format of figure 8 many of those not charted are sufficiently conspicuous in the communiques.

The 10 "probable states"—These are probable discomfort, apprehension, joy, urgency, relief, anger, conflict, sorrow, activation, and effectiveness; this is an extension of an earlier list (ref. 36). The first eight are distinguished (Appendix B) on the situational, stimulus, or input side. Output or response data are examined for compatibility information on situational and background variables and in estimation of intensity. Defining situational changes are specified for each of these probable states in terms of the associated stimulus operations of reinforcement theory (refs. 28 and 41 to 46).

Large, gaping holes would be evident in any analysis restricted to observables. They must be filled by assertions having lower warranty—by inferences from available data drawn by the analyst with varying degrees of subjective probability. Some of these statements will be couched in terms of events and processes presumed to have occurred within the pilot, and warranted more by a knowledge of the culture and specific backgrounds known to have shaped the observable and unobservable behaviors of the particular individual, than by data from the immediate situation.

There is no escaping inferences of this type. The only question is the choice of a conceptual framework for their admission, of which there are many. By one type of approach, sometimes called "phenomenological" (ref. 47), situational and response evidence would be examined along with other data and used for reconstruction of the pilot's "psychological world" in terms of how he perceived, thought, and felt at the moment.

One "reinforcement theory" approach, also widely held, assumes that these inner processes, insofar as they are behavioral, consist of stimulus-response sequences and interactions that function in accordance with the same laws as do the observable sequences on which the principles were established in the first place. These processes are known as "mediational' since they provide a presumptive link between observable stimuli and responses. The greater complexity of human behavior is attributed largely to the great number and variety of intervening and interacting stimulus-response chains evolving from man's capacity to manipulate words and other surrogates for things-in-themselves.

Apart from the comparative theoretical merits of these contending views, the reinforcement-theory approach has practical advantages in this application. Most important is the fact that it reinforces patterns of analysis which make the most of observables before proceeding with reluctance and caution to inferential assertions of lower warranty when they are needed to fill the gaps in making a prediction as to the probable state of the pilot. Moreover, pilots and practically oriented personnel are more likely to accept analyses emphasizing observable situational and response data than one that presumes to know how the pilot is "seeing things," "thinking," and "feeling."

This use of reinforcement theory in combination with Barker's "stream of behavior" concepts and techniques, with their roots in the "phenomenological" approach, is methodologically defensible. The first is concerned exclusively with the search for relationships or "functions" exhibited by all behavior; the latter begins by observing sequences of behavioral events as they occur, in a given spatial, temporal, and social setting, and orders the situational and response data in a manner that can facilitate a "functional analysis."

Assessing the evidence—Three types of judgments are required. The first are "stimulus class-membership judgments." The rater examines the situational changes and stimuli occurring immediately before and during a communique, as portrayed in the "stream of behavior" format (fig. 8), to detect equivalencies between these and the defining situations of the respective probable states. The complexity of these judgments at times is evident in Appendix B.

"Response compatibility judgments" are the second type. The fact that idiosyncratic patterns of expressive and coping behaviors are common among human adults is not surprising in view of differences in genetics and conditioning histories. It is for this reason that probable states are most readily distinguished on the basis of situational evidence (and that generalizing of speech-state relations from one pilot to another without supporting data has been avoided). On the other hand, if response evidence were of no value, it would be ignored. The truth lies somewhere in between. Many joyful, angry, apprehensive, "conflictful," and energetic responses can be identified without situational evidence, particularly if the individuals involved are known. When the sequence of situations in which the responses are embedded is also known, judgments of compatibility are relatively easy. Detection of incompatible responses is the most important part of it.

Raters make intensity estimates also, mapping situational and response evidence onto integer scales with due allowances for the residual and cumulative effects of situations earlier in the sequence and for pilot-selection and training procedures.

The rating sheet has a list of communique numbers in one column, followed by 10 columns each headed by a probable-state name, with each row containing the digit set 123456789. The rater is directed to consider previously identified sets of communiques bracketed by a "situational factor" on the left or an "episode" on the right, to examine the evidence, and to chart first the changes in the states most clearly indicated by the evidence. Other states are then adjusted to allow for the passage of time and for interactions with the new situation.

ALTERNATIVES TO COMPLETE PROBABLE-STATE ANALYSIS

Probable-state analysis (PSA) is tedious and time-consuming. Gathering the evidence and placing it in correspondence with communiques is the most demanding task. But failure in this is a crucial failure.

The approach recommended in the main text reserves complete PSA for the final stage of the validation process, at which time it is performed with the Group as the unit of analysis. The exploratory search for valid speech measures is conducted with simpler criterion techniques applied to simulator and flight situations. Coordinated laboratory studies are employed to tease out relations confounded in the complexity of operational sequences. This section identifies some of these simpler methods.

Restricting the number of ratings—An attempt to substitute two state dimensions—Activation Level and Hedonic Tone—disclosed that weighting and averaging of the intense pleasant and unpleasant aspects of launch, reentry, and some other situations equated them to periods more nearly average in both respects. Splitting of this bipolar scale into two unipolar ones—probable pleasantness and probable unpleasantness—seems desirable.

Restricting the information considered to that in communiques—After some familiarization with the terminology of space flight and with the circumstances of a particular flight, one can rate the state of a pilot over the time span of a Group or communique, using only the transcriptions. However, Davitz (ref. 48) found that speaker-state judgments from spoken sentences are different (presumably better) than those from text alone.

Restricting the analysis to periods of flight clearly evidencing state changes of interest—Emotionally toned incidents can be identified, rated, and classified to represent the clearest contrasts among the several states. Speech-parameter differences are examined for patterns.

Other—The intensive examination (main text) or relations between excursions of a time plot and concurrent events typifies what may be called "coincidence methods"—a term reflecting legitimate interests in co-occurrences and covariation, as well as the inferential hazards of these techniques. The mood ratings and task-performance criteria used to assess responses of Mercury astronauts to stress (ref. 49) offer interesting possibilities. The comparison of flight phases within the framework of Time-Line Analysis is illustrated in the main text. Correlations with heart rate and other physiological measures is another worthwhile direction, particularly if speech measures are computed on a Group basis and if heart rate is averaged over the time span of each Group.

Page intentionally left blank

APPENDIX B

PROBABLE-STATE DEFINITIONS

Interruption of the text is avoided by group citation of most references: refs. 28, 41 to 46, 50, and 51. The following are the easiest sources of the concepts and nomenclature of reinforcement theory: refs. 43, 52, and 53.

Probable discomfort—In one defining situation the individual is exposed to aversive stimulation such as environmental extremes, noxious substances, aches, pains, irritations, or other physiological conditions. In the other he is deprived of such things as food, water, activity, rest, or sex. These two classes of "primary reinforcing stimuli" are termed negative (S^{-R}) and positive (S^{+R}), respectively.

The degree of discomfort or "stress" resulting from the specific exposure or deprivation is estimated from the situational change itself, from the pilot's responses to it, and from general information on the probable time course of its effects. Definitely it is not assumed that all deprivations and all noxious stimulations produce the same "state," but only that the degree of discomfort resulting from the set of those occurring at any moment can be estimated, and that these estimates will correlate with speech-parameter changes.

Probable apprehension—The defining situation is the occurrence of any stimulus which, through prior conditioning, has become a sign of punishing consequences that the individual cannot avoid or terminate. The signal is known as a "negative, secondary reinforcing stimulus" (S^{-r}).

Two types of threatening situations are distinguished when the ratings are made. The first concerns physical dangers and discomforts such as the onset of a serious thruster problem, a landing-bag-deployment signal of unknown origin, a premature 0.05-gravity light, or a substantial rise in capsule temperature after failure of all efforts to control it. The second type contains threats of negative social evaluation or self-evaluation. Astronauts have been selected from a culture, by a process, and for a task, each of which places a premium on achievement and success. One pilot stated forthrightly, "I thought this was a chance for immortality"—the ultimate in social approval.

> Central in their personalities are strong needs for achievement and mastery. These are men who must do, and do well, and this quality is obvious early in their histories. . .. By the same token, their potential for disappointment is high. With strong needs for achievement and strong feelings of individual responsibility, failure can be disturbing, for it cannot readily be minimized or rationalized. . .. As committed men, disappointments are keenly felt; as ego-strong men, hope is sustained and disappointment leads to renewed effort. Similarly, effects are readily aroused and strongly felt, but there is good control of potentially disabling effects on behavior (ref. 49).

A flight surgeon said of one pilot that he was more concerned about efficient performance than about external dangers. For these reasons anything that threatens or appears to threaten the full and competent achievement of both major and minor flight goals is interpreted as situational evidence of some degree of "probable apprehension" even if the only consequence is disapproval (S^{-gr}) (or withholding of approval) by associates. Situations and performances likely to generate self-disapprovals are also in this category. Negative self-evaluations become fairly reliable indicators of probable social disapprovals, and, by virtue of their occurrence over a wide range of situations and consequences, they become "generalized" and functions as S^{-gr}'s in their own right.

The common element that gives this category its theoretical integrity is the S^{-r} signal, S^{-gr} being a subset. This reduction enhances confidence that ratings, based on situational and response evidence of such apparent diversity, may reflect a unitary state.

Probable joy—The defining situation is the occurrence of any stimulus that, through prior conditioning, has become a sign (S^{+r}) of probable, positive reinforcement—of the occurrence of attainment of an S^{+R}; or of the cessation, termination, or avoidance of an S^{-R} or an antecedent S^{-r}.

The two types of situational evidence considered parallel exactly those of "probable apprehension." The first type includes an acceleration profile duplicating the expected one experienced in simulation; onset of a green retroattitude light, the appearance of the chutes, and a message from a recovery plane that "I have you visually." These S^{+r}'s are obviously related in the operational system to primary positive reinforcements. The acceleration example illustrates the fact that some strong S^{-R}'s can function simultaneously as S^{+r}'s. Other examples include the temporary envelopment of the capsule in flames as booster engines drop away, which would be terrifying had not photographic evidence established it as verification (S^{+r}) of proper staging. The "fireball" during reentry ordinarily signifies (S^{+r}) proper functioning of the ablative heat shield. The S^{-r} components of these situations probably still evoke some "probable apprehension." Allowance must be made for these complexities in rating of all three states discussed so far.

In situations of the second type the S^{+r}, regardless of origin, signals positive social evaluations (S^{+or}'s). When the pilot has the requisite feedback, the mere occurrence of a successful or outstanding performance may be taken as presumptive evidence of the occurrence of favorable self-evaluative responses, and of the joy occasioned by these as predictors of social approvals—if one assumes that the attainment is a significant one.

In one flight the achievement of control over suit temperature by following a carefully developed plan is an illustration. As evidence (S^{+r}'s) of successful control accumulated, there was justifiable "pride and joy" in the accomplishment, followed by approvals from the ground (S^{+or}'s). In this case, primary reinforcements (S^{+R}'s) also resulted (here without social mediation) from control over suit temperature for the rest of the flight. (The joyful exhilaration on attainment of orbit, and its intensification by other factors, is discussed in the main text.)

Probable urgency—This is an estimate of task pressures. The high-urgency situation includes both S^{-r} and S^{+r} components of a special class. The S^{-r} components signal the highly probable onset, at the end of a relatively fixed period, of either a severely punishing consequence or an immutable sequence culminating in one. The S^{+r} components signal concurrently the probability that the required performances can be discovered if necessary and executed in time for successful avoidance of the undesirable consequences. Siegel and Wolf (ref. 54) suggest the ratio of the time required for completion of essential tasks, to the time available, as the appropriate index.

When the S^{-r} elements justify a high "probable urgency" rating, they also increase "probable apprehension" and "probable activation" and could be absorbed by them. As a separate item, the "probable emergency" charts one important class of tension-producing situations.

Probable relief—These situations are identified by substantial reduction or termination of S^{-R} or S^{-r} stimulation. The combination of "relief and joy," encountered for example as launch and reentry sequences near completion, is due to the fact that the reductions and terminations occasioning the relief also signify (S^{+r}) successful completion of the phase or mission. When the pilot has contributed to the situational change, self-approval or social approval enters the picture, as in the example of suit-temperature control.

Probable anger—The interruption of a stimulus-response sequence that has customarily resulted in positive reinforcement (either procurement of positive reinforcements or escape from, termination of, or reduction of negative reinforcements) is the defining situational input. The breaking of sequences leading to social approval or self-approval is a special case.

The holds and delays of countdown are an example, with one pilot providing substantiating evidence on the response side. During one of many long holds he is reported to have gone on the intercom with "his only terse remark of the day": "I'm cooler than you are. Why don't you fix your little problem and light this candle:" Pressure suits and other physical constraints qualify by interfering with normal responses for relieving discomforts and procuring positive reinforcements. Requesting an explanation and not getting it in the customary manner is another.

Probable conflict—Conflict situations present stimuli requiring two or more incompatible responses or response sequences. Conflict between the attractiveness of the view (or of other phenomena) and the programmed task requirements is one example. Attempts at self-control also evidence conflict. They are responses to situations containing two types of reinforcement contingencies: one evoking the angry, anxious, dozing, or other responses to be controlled; the other, the controlling responses. Successful controlling responses generate stimulus-response sequences that are incompatible with the to-be-controlled responses; for example, one pilot calmed himself during a hold by the self-command, "You're building up too fast. Slow down. Relax."

Probable sorrow—The situational change involves irretrievable loss of the S^{+r}'s that have in the past indicated a high degree of probability that one's responses would be positively reinforced. When the opportunity to achieve a highly desirable flight objective is irretrievably lost, the states that follow are likely to include "probable anger" when the scheduled sequence is interrupted, some "probable apprehension" if a hazard or social disapproval is indicated, and finally "probable sorrow" (regret) when the permanence of the loss is recognized.

The sadness that alternates with anxiety when death is inevitable, and death's approach is slow enough to permit such alternations, also fits the definition. The S^{+R}'s about to be lost involve stimuli generated in one's own body, including those stimulus-response sequences called the "ego" or "self." To the extent that the situation is highly aversive, anticipation of "probable relief" is an additional component.

Probable activation—This is the "arousal syndrome." It is estimated mainly from response (output) information, with heart rate serving as an available index for most pilots. Other evidence considered includes the range of stimuli to which responses are being made at a given time; the latency, rate, and adequacy of the responses; introspective or retrospective reports of elation, drowsiness, or not being "on top of things"; and the text of communiques. Situational evidence is examined to ensure compatibility with response evidence and to support intensity estimates; this procedure reverses that of the eight situationally defined states.

Probable effectiveness—This is an overall assessment of the pilot's ability to process information and perform required tasks; it is based on all the evidence, whether included in the preceding ratings or not. Job-performance information is given most weight: for example, switch-throwing errors; double-authority errors in attitude-control; time taken to detect an error; looking for a constellation where it would have been if the flight had not been delayed; and pursuing lower-priority tasks to the detriment of higher-priority ones. Introspective reports also are relevant. "Probable confusion," deleted from the preproposal suggestions, is absorbed by the information-processing aspects of "probable effectiveness," and by the situational aspects of "probable urgency." Boredom and other states not specifically listed are similarly encompassed.

Page intentionally left blank

REFERENCES

1. BARKER, R. G.; AND WRIGHT, H. F.: Midwest and Its Children. Harper and Row Publishers, Inc., 1955.
2. BARKER, R. G.; WRIGHT, H. F.; BARKER, L. S.; AND SCHOGGEN, M.: Specimen Records of American and English Children. University of Kansas Publications, 1961.
3. BARKER, R. G., ED.: The Stream of Behavior. Appleton-Century-Crofts, Inc., 1963.
4. STARKWEATHER, J. A.: Measurable Dimensions of Vocal Behavior. Presented at meeting of American Association for the Advancement of Science (Denver, Colo.), Dec. 1961.
5. POPOV, V. A.; SIMONOV, P. V.; TISHCHENKO, A. G.; FROLOV, M. V.; AND KHACHATUR'YANTS, L. S.: Analysis of International Characteristics of Speech as a Criterion of the Emotional State of Man under Conditions of Space Flight. Zh. Vysshei Nervnoi Deyatel'nosti im. I. P. Pavlova, vol. 16, no. 6, 1966, pp. 974–983. (Translation TT 67-30554 available from CFSTI.)
6. VERZEANO, M.; AND FINESINGER, J. E.: An Automatic Analyzer for the Study of Speech in Interaction and in Free Association. Science, vol. 109, 1949, pp. 45–46.
7. VERZEANO, M.: Time-Patterns of Speech in Normal Subjects. J. Speech Hearing Disorders, vol. 15, 1950, pp. 197–201.
8. VERZEANO, M.: Time-Patterns of Speech in Normal Subjects. J. Speech Hearing Disorders, vol. 16, 1951, pp. 346–350.
9. BRADY, P. T.: A Technique for Investigating On-Off Patterns of Speech. Bell System Tech. J., vol. 44, 1965, pp. 1–22.
10. GOLDMAN-EISLER, F.: The Distribution of Pause Durations in Speech. Langu. Speech., vol. 4, 1961, pp. 232–239.
11. GOLDMAN-EISLER, F.: Speech-Breathing Activity—A Measure of Tension and Effect During Interviews. Brit. J. Psychol., vol. 46, 1955, pp. 53–63.
12. GOLDMAN-EISLER, F.: Speech-Breathing Activity and Content in Psychiatric Interviews. Brit. J. Med. Psychol., vol. 29, 1956, pp. 35–48.
13. GOLDMAN-EISLER, F.: The Determinants of Rate of Speech Output and Their Mutual Relations. J. Psychosomat. Res., vol. 1, 1956, pp. 137–143.
14. HENDERSON, A.; GOLDMAN-EISLER, F.; AND SKARBEK, A.: Temporal Patterns of Cognitive Activity and Breath Control in Speech. Langu. Speech, vol. 8, 1965, pp. 236–242.
15. HENDERSON, A.; GOLDMAN-EISLER, F.; AND SKARBEK, A.: Sequential Temporal Patterns in Spontaneous Speech. Langu. Speech, vol. 9, 1966, pp. 207–216.
16. GOLDMAN-Eisler, F.: The Predictability of Words in Contexts. Quart. J. Exp. Psychol., vol. 10, 1958, pp. 96–106.
17. GOLDMAN-EISLER, F.: Significance of Changes in the Rate of Articulation. Langu. Speech, vol. 4, 1961, pp. 171–174.
18. FONARGY, I.; AND MAGDICS, K.: Speed of Utterance in Phases of Different Lengths. Langu. Speech, vol. 3, 1960, pp. 179–192.
19. STARKWEATHER, J. A.: A Speech Rate Meter for Vocal Behavior Analysis. J. Exp. Ann. Behav., vol. 3, no. 2, 1960, pp. 111–114.
20. NEWCOMB, W. B.: The Perceptual Basis of Syllable Boundary. Presented at Summer Meeting, Linguistic Society of America (Austin, Texas), 1963. (Available from Electronics Division, General Dynamics Corp., Rochester, N.Y.)
21. HARRIS, J. S.: Voice Communication: Effect of Stress Conditions on Speaker Intelligibility. Contract N6ori-104, Purdue University, 1948.
22. FAIRBANKS, G.; AND PRONOVOST, W.: An Experimental Study of the Pitch Characteristics of the Voice during the Expression of Emotion. Speech Monogr., vol. 6, no. 1, 1939, pp. 87–104.
23. HUTTAR, G. L.: Some Relations Between Emotions and the Prosodic Parameters of Speech. SCRL Monograph No. 1, 1967.
24. COCHRAN, W. G.; AND COX, G. M.: Experimental Designs. John Wiley and Sons, Inc., 1950.
25. ESTES, W.; AND SKINNER, B. F.: Some Quantitative Properties of Anxiety. J. Exp. Psychol., vol. 29, 1941, pp 390–400.
26. NEWCOMB, W. B.: Some Tempo Manifestations of the Terminals of English. University Microfilms, Inc., 1960.
27. SCHROEDER, M. R.: Vocoders: Analysis and Synthesis of Speech. Proc. IEEE, vol. 54, no. 5, 1966, pp. 720–734.
28. SKINNER, B. F.: Verbal Behavior. Appleton-Century-Crofts, Inc., 1957.
29. GREEN, D. M.: Consistency of Auditory Detection Judgments. Psychol. Rev., vol. 71, 1964, pp. 392–407.
30. SENDERS, V. L.; AND SOWARD, A.: Analysis of Response Frequencies in the Setting of a Psychophysical Experiment. Amer. J. Psychol., vol. 65, 1952, pp. 358–374.
31. SURGENT, L. V.: The Speech-Analysis and Speaker-State Study Capability of the Communications Sciences Laboratory. Electronics Division, General Dynamics Corp., Rochester, N.Y., Feb. 1965.
32. SURGENT, L. V.: Acoustical Analysis of Astronaut Speech. Vol. I (main text). Vol. II Appendices. Electronics Division, General Dynamics Corp., Rochester, N.Y., Feb. 1967 (NASw-1252).

33. HEFFERLINE, R. F.: An Experimental Study of Avoidance. Genet. Psychol. Monogr., vol. 42, 1950, pp. 231–334.
34. SIDMAN, M.: Two Temporal Parameters of the Maintenance of Avoidance Behavior by the White Rat. J. Comp. Physiol. Psychol., vol. 46, 1953, pp. 253–261.
35. SURGENT, L. V.: Some Directions in Speaker-State Research. Electronics Division, General Dynamics Corp., Rochester, N.Y., Aug. 1968 (Report No. 68–126).
36. SURGENT, L. V.: The Applicability of Acoustical Speech Analysis Techniques to the Monitoring of Spacecraft Crews: A Preprosal Submittal. Electronics Division, General Dynamics Corp., 1960, Rochester, N.Y., June 1964.
37. SOSKIN, W. F.; AND JOHN, V. P.: The Study of Spontaneous Talk. The Stream of Behavior, R. G. Barker, ed., Appleton-Century-Crofts, Inc., 1963.
38. WOODWORTH, R. S.: Situation- and Goal-Set. Amer. J. Psychol., vol. 50, 1937, pp. 130–140.
39. SCHOGGEN, P.: Environmental Forces in the Everyday Lives of Children. The Stream of Behavior, R. G. Barker, ed., Appleton-Century-Crofts, Inc., 1963.
40. SKINNER, B. F.: The Behavior of Organisms. Appleton-Century-Crofts, Inc., 1938.
41. HUNT, W. A.: Recent Developments in the Field of Emotion. Psychol. Bull., vol. 38, 1941, pp. 249–276.
42. HUNT, J. MCV.; COLE, M. L.; AND REIS, E. E. S.: Situational Cues Distinguishing Anger, Fear and Sorrow. Amer. J. Psychol., vol. 71, 1958, pp. 136–151.
43. KELLER, F. S.; AND SCHOENFELD, W. N.: Principles of Psychology, Appleton-Century-Crofts, Inc., 1950.
44. MOWRER, O. H.: Learning Theory and Behavior. John Wiley and Sons, Inc., 1960.
45. MOWRER, O. H.: Learning Theory and the Symbolic Processes. John Wiley and Sons, Inc., 1960.
46. SKINNER, B. F.: Science and Human Behavior. The MacMillan Co., 1953.
47. WANN, T. W., ED.: Behaviorism and Phenomenology. Contrasting Bases for Modern Psychology, University of Chicago Press, 1964.
48. DAVITZ, J. R., ED.: The Communication of Emotional Meaning. McGraw-Hill Book Co., 1964.
49. KORCHIN, S. J.; AND RUFF, G. E.: Personality Characteristics of the Mercury Astronauts. The Threat of Impending Disaster. Contributions to the Psychology of Stress, G. H. Grosser; H. Wechsler; and M. Greenblatt, eds., M.I.T. Press, 1964.
50. MCCLELLAND, D. C.; ATKINSON, J. W.; CLARK, R. A.; AND LOWELL, E. L.: The Achievement Motive. Appleton-Century-Crofts, Inc., 1953.
51. MCCLELLAND, D. C.: Notes for a Revised Theory of Motivation. Studies in Motivation, D. C. McClelland, ed., Appleton-Century-Crofts, Inc., 1955.
52. HOLLAND, J. G.; AND SKINNER, B. F.: The Analysis of Behavior. McGraw-Hill Book Co., 1961.
53. ULRICH, R.; STACHNIK, T.; AND MABRY, J.: Control of Human Behavior. Scott, Foresman and Co., 1966.
54. SIEGEL, A. T.; AND WOLF, J. J.: Computer Simulation of Man-Machine Systems. Unusual Environments and Human Behavior, N. M. Burns; R. M. Chambers; and E. Hendler, eds., The Free Press of Glencoe, 1963.

☆ U.S. GOVERNMENT PRINTING OFFICE: 1971 0–412–732