NASA CONTRACTOR REPORT

NASA CR-1720

NASA CR-17
c.1

TECH LIBRARY KAFB, NM
0060279

LOAN COPY: RETURN TO
AFWL (DOGL)
KIRTLAND AFB, N. M.

INTRODUCTION TO A BIOLOGICAL SYSTEMS SCIENCE

by E. Bloch, S. Cardon, A. Iberall, D. Jacobowitz, K. Kornacker, L. Lipetz, W. McCulloch, J. Urquhart, M. Weinberg, and F. Yates

Prepared by
GENERAL TECHNICAL SERVICES, INC.
for
U. S. ARMY RESEARCH OFFICE
and

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION • WASHINGTON, D. C. • FEBRUARY 1971

TECH LIBRARY KAFB, NM

0060779

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
NASA CR-1720		

4. Title and Subtitle	5. Report Date
INTRODUCTION TO A BIOLOGICAL SYSTEMS SCIENCE	March 1971
	6. Performing Organization Code
7. Author(s) See Supplementary Notes (item 15).	8. Performing Organization Report No.
9. Performing Organization Name and Address General Technical Services, Inc. Upper Darby, Pennsylvania 19082	10. Work Unit No.
	11. Contract or Grant No. NASw-1815
12. Sponsoring Agency Name and Address National Aeronautics & Space Administration Washington, D. C. 20546	13. Type of Report and Period Covered Contractor Report
	14. Sponsoring Agency Code

15. Supplementary Notes The authors are as follows: E. Bloch (Case-WRU), S. Cardon (GTS), A. Iberall (GTS), D. Jacobowitz (U.Pa.), K. Kornacker (OSU), L. Lipetz (OSU), W. McCulloch (MIT), J. Urquhart (USC), M. Weinberg (GTS), and F. Yates (USC)

16. Abstract

An attempt has been made to provide an introduction to a biological systems science by the following strategy: A set of principles are put forth as a general systems' science theoretic. They include the idea that a complex system is caught up in a large number of oscillatory chains - thermodynamic 'engine' cycles - which maintain the average regulated state; that an autonomously operating system is marginally unstable so that it cannot come to rest; that such a system operates within many modes; that the autonomous system has to develop a complex behavior which is ordered in temporal and spatial scale via these modes.

Various of the biological subsystems are then discussed and analyzed, in nearly unitary fashion, within the overall systems' paradigm. The subsystems treated include the membrane, the microvasculature, the nervous autonomic system, the cardiovascular system, the biochemical chains, the higher nervous structure and function, and the hormonal foundations for behavior. The integration of behavior was intended, but not completed.

17. Key Words (Selected by Author(s))		18. Distribution Statement
biology	behavior	Unclassified - Unlimited
systems	biophysics	
cybernetics	homeokinesis	
nervous system	homeostasis	
endocrine system	brain	

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price*
Unclassified	Unclassified	335	$3.00

*For sale by the National Technical Information Service, Springfield, Virginia 22151

FOREWORD

On behalf of the International Union of Biological Science (IUBS), as a focus of international organizations in all the branches of biology, C. H. Waddington undertook to examine the question[1] whether a theoretical biology might be developed, as a disciplinary science, in some sense similar to theoretical physics. It has been said that the IUBS felt that it was its timely duty to begin such an inquiry. We collaborators, joined to give each other counsel, have responded to that challenge. We hope that our deliberations have borne one piece of fruit, an introduction to a biosystems science.

However, we admonish any biological reader, should he expect to find a unitary description for biology, to at least scan some of the classics in theoretical physics.[2] In our opinion he will find in that literature a structure built around many themes, and based on a complex pluralistic metaphysics, rather than any unitary structure. Awareness of the many unresolved fundamental questions in biology may be equally perplexing.

DEDICATION

Our dear friend Warren McCulloch died during the preparation of this work. The loss to us and to our expectation that he would help us bring the concrete questions about behavior into focus is very great. We grieve his passing, and we dedicate this work to his spirit.

[1] C. H. Waddington "Towards a Theoretical Biology", 1. Prolegomena (1968); 2. Sketches (1969); 3. Drafts (in press); Aldine.

[2] See such textbooks as Page, Joos, or Landau and Lifshitz.

TABLE OF CONTENTS

Page

A. A PHYSICAL INTRODUCTION..... 1

I. PRELIMINARIES..... 1

1. On a Paradigm for a Biological Systems Science..... 1

2. Definition of a System - the Isomorphic Ideal..... 2

3. Systems Science and Other Sciences..... 2

4. Abstract Transformations in Physical Systems..... 3

5. A Note on the Theory of Large Versus Small Systems..... 6

6. Some General Concepts for Large Systems Analysis..... 6

7. 'Directed' Chains - Characterizing a Large System..... 8

8. A Statistical-Mechanical View of Life and Life's'Purposes'.. 11

9. On the Levels of Biosystem Science..... 14

10. The Normal Operating State of a Biosystem..... 15

11. Defining the Normal Response of a Biosystem..... 18

12. The Phases of a System - Start Up, Life, Degradation..... 21

13. The Point of View of Systems Design..... 22

II. THE ORGANIZING PRINCIPLE OF COMPLEX LIVING SYSTEMS..... 24

1. Instability and the Nature of Life..... 24

2. The Spectroscopy of Man - Organization in Space and Time.... 26

3. The High-Frequency Biochemical and Electrical Foundations... 28

4. The Medium-Frequency Range of Behavior..... 32

5. The Low-Frequency Range - Life's Foci..... 34

Page

III. HIERARCHICAL REGULATION IN THE COMPLEX BIOLOGICAL ORGANISM..... 37

1. Biophysical Preliminaries..... 37

2. Dynamic Limit Cycles in Mammals - The Foundation for a Belief in Homeokinesis..... 39

3. Describing Homeokinesis - Dynamic Regulation of the Organism's Interior..... 39

4. Hierarchical Regulation..... 40

5. Applying the Concept of Hierarchical Regulation..... 43

IV. SUMMARY OF SOME OF THE ABSTRACT PRINCIPLES FOR A BIOSYSTEMS SCIENCE..... 50

Bibliography..... 53

B. INVESTIGATING SOME OF THE MAJOR BIOSYSTEMS 56

V. AN INTRODUCTION TO STRUCTURES AT THE LOWER LEVELS OF ORGANIZATION..... 56

1. The Cell Membrane - Composition and Physical Characteristics..... 56

2. Informational and Functional Substances in Membranes..... 60

3. Theories of Membrane Structure..... 62

4. Other Concepts of Membrane Structure..... 62

5. Electron Transport Membrane Systems..... 63

6. Ion Pumps in Membranes..... 64

7. Resting Membrane Potentials..... 64

8. Action Potentials..... 65

9. Coupling Between Solute and Water Transport..... 65

10. Other Ion Pumps..... 65

Page

11. Transport Modes and Mechanisms 66

12. Some Overall Remarks 68

13. Summary 69

14. Addendum - A Speculative Note on a Dynamical Hierarchy of Active Transport Systems 69

Bibliography 73

VI. SOME COMMENTS ON CONTROL OF THE MICROVASCULAR SYSTEM 76

1. Introduction 76

2. Background 76

3. The Organizational Level of Microvascular Control 84

4. Addendum I - Microvascular Geometry 98

5. Addendum II - Transcapillary Exchange 102

Bibliography 108

VII. THOUGHTS CONCERNING AUTONOMIC NERVOUS SYSTEMS CONTROLLING INFLUENCES 113

1. The Autonomic Ganglia 113

2. Autonomic Action and Interactions at the End Organ 115

3. The Cardiovascular System 120

Bibliography 128

VIII. THE CARDIOVASCULAR SYSTEM 130

1. Introduction 130

2. Temporal Fractioning of Cardiovascular-Metabolic Phenomena .. 149

3. Rationale for Focussing Initial Discussion on the 400 Second Region of the CV Spectrum 151

Page

4. Some Elementary Parameters in the Cardiovascular System..... 155

5. The Range of Values of the Elementary Parameters in Rest and Exercise..... 156

6. An Engineering Definition of Systems Relations Governing Cardiac Output..... 165

7. A Preliminary View of the Microcirculation in Skeletal Muscle..... 192

Bibliography..... 195

IX. A PHYSICAL VIEW OF METABOLIC SYSTEMS - THE BIOCHEMICAL TRANSITION FROM ORGAN SYSTEMS TO THE HIGHER CENTRAL NERVOUS SYSTEM..... 198

1. General Principles of Metabolic Systems..... 198

2. Life Without Nervous Systems..... 211

3. Time Domain of Cellular Processes..... 217

4. The Structural Level Defining a Metabolic State..... 219

5. Bioenergetics in Animals at the Organ System Level..... 221

6. Concluding Comments about Bioenergetic Systems of Animals... 223

Bibliography..... 229

X. THE ROLE OF THE NERVOUS SYSTEM IN THE ORGANIZATION AND INTEGRATION OF LIVING SYSTEMS..... 230

1. Cellular Dynamics..... 230

2. Intercellular Communication..... 231

3. Information Input to the Nervous System..... 236

4. Transfer of Information Along a Nerve Cell..... 238

5. Encoders..... 240

Page

6. Categorization of Input Information - Pattern Recognition... 247

7. Evaluation of Input Information - 'Subgoals' as Hierarchical Entities..... 257

8. Output Systems..... 261

9. Coordination of the Neural and Hormonal Systems..... 265

10. Selection of Action Routines for Attaining a Subgoal..... 269

11. Information Flow in the Nervous System..... 271

12. The Ontogeny of Behavior..... 273

Bibliography..... 284

XI. THE CHEMICAL ORIGINS OF MAMMALIAN BEHAVIOR..... 294

1. Periodic Modal Behavior..... 294

2. Chemical Communication Pathways To and From Brain..... 300

3. Thermodynamics and Behavior..... 301

4. Long-Term Periodicities of Human Behavior..... 303

5. The Full Spectrum for the Life Phase of Adults..... 304

6. Hormones and Behavior - Sexual Behavior..... 307

7. Hormones, Metabolites and Behavior - Summary..... 308

8. Summary of the Physical View of Behavior..... 309

Bibliography..... 316

XII. SUMMARY, OVERVIEW, AND DISCUSSION..... 319

XIII. ACKNOWLEDGEMENT..... 327

A. A PHYSICAL INTRODUCTION

I. PRELIMINARIES

1. On the Need for a Paradigm for a Biological Systems Science

The contributors to this study, although motivated by a common will to get a task done, found impediments to progress that arose time and again, not so much at the beginning of the effort, but as the conflicts in point of view forced a greater degree of inward searching upon each participant. Such impediments usually arise among workers from different disciplines whenever they tackle an interdisciplinary problem, because they operate from different philosophic frames of reference; they hold different metaphysics -- different views of the operating characteristics of science. They literally do not, at first, understand each other. In the language emphasized by Kuhn (1), they operate with different 'paradigms'. This difficulty must be kept in mind by the reader, lest our attempt at a biological systems science appear grandiose or hopeless to readers who fail to grasp the awkwardness of interdisciplinary efforts. We have tried to overcome the obstacles.

The current and past generations of biologists have been steeped in an experimental tradition. Biology consists of intensive experimental exploration, observation, and measurement. General principles or extrapolations beyond the range of measurement and observation are suspect. The cause of the reluctance to abstract is historical; an earlier era was too bold and free wheeling, and encouraged incompetents (as well as competents) to hold forth without evidence. Yet what is underestimated by today's scepticism is the value of this freely innovative style in which physicist and chemist and mathematician were often partners with the biologist and physician in the earlier era of biological discovery.

Fortunately, there is a group of physical scientists and engineers who are seeking new scientific fields to explore with a rather freer scientific methodology, and a group of biologists who welcome the fresh approach. Together they search for broad and abstract principles and summaries. Characterizing their efforts by a phrase which is popular among them today, they seek to 'model the system'. The question that may be asked is whether they are willing to 'capitulate to biology', to learn and absorb the real properties of the living system.

In seeking to model systems, there are two kinds of models. One kind is a 'formal' model, in which data from experiments are represented by relations of convenience (e.g., mathematical curve fitting), or better by quasi-physical intuitions. The other is an 'isomorphic' model, in which firm physical reasoning suggests relations that should agree both in form and function with experimental measurements. The closer we can come to isomorphic models, the better chance we have of creating a biological systems science. Thus our basic question is: to what extent is a physically-based isomorphism possible in biology? To answer, we adopt a common physical paradigm for our work, to be explained below, with the certainty that we do not yet know the answer to the basic question. The answer must come, however, from efforts such as this.

2. Definition of a System - the Isomorphic Ideal

A system is described in two contexts. First, a system is an actual, localized physical structure, embedded in the objective world, undergoing transformation and change. Second, it is a logical structure, describable by a metalanguage (which is not here defined[1]). That is, it may be a set of objects and relations (which can exist in some world, preferably an objective world which is publically available), a set of definitions, a set of postulates, and a closed or open set of axioms.

It is often the aim of science to attempt to make the two descriptions isomorphic, so that each description is 'essentially' capable of representing in one to one fashion all of the states, both static and dynamic, that the two system descriptions can take on.[2] Such an aim is too ambitious, because natural systems can in principle pass through vast domains of space and time. Therefore, we accept a more limited scope for isomorphism - namely that the states of the logical and physical systems should correspond within a prescribed limited domain of space and time. For two systems to show 'essential' correspondence or to be 'essentially' represented requires that the systems have the same distribution and types of singular states of motion. Then, even though there may be some local, noncorresponding space and time distortions in their behavior, these do not affect them essentially, and they remain 'isomorphic' in an operational sense.

3. Systems Science and Other Sciences

Systems science - and we are interested in the particular case of biological systems science - is not coextensive with physics, chemistry, or biology. It is certainly not coextensive in the case of large or complex systems. If it were, there would be no need for systems science, for it would be the natural sciences. Our problem is to discover whether we can build an isomorphic systems science for the (large) biological system. Obviously any fertile mammalian mother (given a fertile father) can build[3] a biological system. Can we represent the system, either conceptually or in the laboratory?

[1] Except to state that it is the language used within the common culture shared by those involved in examining a system.

[2] There are other logical systems- paper systems - which may interest the logician, which have nothing to do - as yet - with known physical systems. These are not our concern.

[3] Isomorphism is not restricted to a relation between a physical and a logical system. Two systems each isomorphic to a third system are isomorphic to each other. Thus two real systems can be isomorphic. In this context we would regard the building (self-replicating) process accomplished by parents as isomorphic analogues.

4. Abstract Transformations in Physical Systems

As a first step toward constructing a biological systems science, we may borrow and abstract some useful ideas from physical systems.

First, physical systems have some identifiable form (structure) and function in space and time - namely, their form or structure is localized in space, there are some geometric relations between entities that may be followed in space and time, and there are certain transformations among parts that can take place.

We can loosely classify the transformations, by physical analogue.

(a) Transfer operations and apparatus. - There are transfer operations - involving fixed formal constraints in a physical field - that entities entering the field will undergo. A common example is geometric transformations that may take place at the many stations of a complex assembly line.

The concept of transfer operations came into existence in mechanical engineering practice to represent those operations and apparatus which transform the physical state of components in space and time.[1]

In approaching the subject of automatic control, the Russians formally introduced these ideas (see for examples (2)). In their view a coin-changing machine, a textile weaving machine, a production line, shift registers, coded templates, and pinball machines could all be regarded as examples of transfer apparatus. What seemed strange to American automatic control thinking was that a transfer apparatus appeared to be merely a static structure. Yet, if an energetic entity is inserted in the field, space-time transformations result, and the entity follows a trajectory or guided path.

There are two kinds of constraints to which the guided paths offered by the transfer apparatus may be subject - one is the holonomic or integrable constraint, whereby the transformation in space and time is fully describable, and fully determinate. The second - the nonholonomic constraint - is not expressible by differential relations. The most interesting case, for biological systems, is the one in which the result of a local transformation has some indeterminate character. Since all natural processes are lossy (entropy is lost to the universe), no transformation is fully determinate. Therefore the important distinction regarding nonholonomically constrained systems concerns how long coherence can be followed compared to the (long term) space and time interval of closure. A dice game in which one

[1] The description by hierarchical levels in a system begins here. We have already noted that descriptions of systems are to be limited in space and time. The idea of levels - temporal 'epochs' or spatial 'layers' of structural organization - then comes into prominence. If we speak at the molecular biological level, the concept of physical state will include chemical association. At macroscopic levels, it will deal only with macro or organized forms.

invariably loses at every toss of the die is quite deterministically fixed, whereas if the house takes only a 'normal' 10% edge one may regard the game as 'fair', or stochastic. In between, there are all kinds of arrangements that the 'designer' or 'fixer' can introduce to modify the coherence.

A biological organism may be regarded as a factory embedded in a complex milieu. Together with its milieu, the combination may be regarded as the biological system. In one view it may be seen as a complex transfer apparatus - like a complex pinball machine: it has fixed stations[1] at which both holonomic and nonholonomic interactions take place, and which exhibit time dependent and space dependent properties that we will attempt to define.[2] A biological system may also be viewed as a complex musical organ with many pipes and many stops. The master player and the winds of the milieu both[3] may create the many musical patterns that represent the external (motor) states of the system and the internal (secretory and flux-like) states of the system.[4]

(b) The chains in man. - While we will speak of the biological organism, loosely meaning living forms from one celled creatures to man, it is man himself that we shall constantly seek out, as our major interest. More generally we will mean mammals, still more generally, we will mean animals, and only in extremes will we mean all of life. Although the biological organism, as a system, actually consists always of a living form and its external environment, we consider at first the organization of the internal environments.

We are concerned with internal chains of causality, e.g., how internal element A is linked in space and time to B, to C, to D, etc. More narrowly, 'chains' also means physical linkage of elements. The resulting arrangement is a system strung out in space. Thus, we prefer not to use a term like 'network' because it tends to imply a more rigid lumped character than we

[1] For example, we can stimulate a nerve and obtain a response time and time again.

[2] Mass fluxes of food, water, etc.; and energy fluxes of sensory inputs come in, and are shuffled around and through the system.

[3] While a literary characterization, we will tentatively identify these as the cortical algorithm that has developed in the CNS, and the patterned waves of endocrines and neuroendocrines and electrical signals that sweep through the brain.

[4] On the other hand, from the point of view of the internal organs, they enter in, on cue in their actions, to help orchestrate the complex patterns of biochemical and bioelectric fluxes that they tend to mediate.

desire.[1] What interests us are those chains that dominate the disposition of informational and power fluxes in space and time domains, especially those in which fluxes are continuous[2].

We recognize five types of chains:

(1.) Degradative or dissipative chains. - Suppose that a flux is found to exist at some time. At some later times - within the realm of interest, the flux may have diminished and ultimately stopped. Descriptively, the system is represented as a transient nonconservative system.

(2.) Lossy maintained chain. - Suppose that a constant flux is maintained. Since all natural processes are ultimately degradative, in order for a flux to be sustained, power must be drawn from some external source. This power may be obtained either by intermittent or continuous energy intake. Such phenomena are viewed in engineering as the theory of D.C. networks, or the theory of steady state networks. The system may be described as a nonconservatively maintained steady state.[3]

(3.) 'Explosive' or 'departing' chain. - There are some chains which leave the (agreed upon) field of observation in time. They depart or explode. After their passage, we can no longer be concerned with them. Their subsequent behavior becomes purely speculative. If, on the other hand, they return at intervals, and are detectable, then they are no longer specified as 'departing' chains, but as repeating chains.

(4.) Periodic or repeating chains. - These chains tend to remain within our observed field, but may depart and return periodically. In general, they are nonlinear. Their performance can be described for the stable steady states of nonlinear nonconservative dissipative systems, by the limit cycles of Poincaré.

(5.) Aperiodic or intermittent chains. - We have no useful theory to account for a universe, nova, sun, or volcano that may blow up once, or at such rare or aperiodic times that regularity cannot be detected (within the time of

[1] Drawing from a number of physical fields, we can think of the physicist's free body, the electrical engineer's network, the chemist's compartment analysis, or the chemical engineer's unit process. Each of these are isolating entities or concepts that aid in analyzing a causal or connected sequence of parts or processes.

[2] Or at least so frequent that we may regard them as continuous. Biologists are familiar with squirt systems, in which fluids are delivered as a sequence of pulses. One must note that we have already taken care of this possibility by suggesting that time-of-action can be bounded from below. Thus, there is a limit at which there is no essential difference between the continuous and the discrete process.

[3] We cannot deal with a systems science of conservative systems, nor are we called on to do so for biological systems, which are open systems. The basic weakness that we face, for conservative systems, is that we cannot discuss their stability against disturbances. For example, even though they can exhibit autonomous periodic phenomena, we cannot project the effect of perturbations.

interest), unless the sources of decay and explosion are known. However, in time and with study, appropriate laws may be discovered for such events.

5. A Note on the Theory of Large Versus Small Systems

It is desirable that a distinction be made between small and large systems.

A small system consists of a few chains, and generally only of one or two (not more than a few) hierarchical levels. Much of the science of physical networks applies only to small systems (as found in mechanical 'n degree of freedom' systems, in electrical and magnetic network problems, in optical field design, in elastic field problems, etc.). For such problems, few or no significant ideas can be added by systems science. Systems science is here completely coextensive with the physical sciences.

On the other hand, large systems have many network chains which occur at more than a few hierarchical levels. Phenomena over a broad frequency spectrum will be involved; there will be many groups of independent causality in these spectra (i.e., each phenomenon may have its own partial spectrum); phenomena will be widely distributed spatially at different levels of organization; the couplings may be poorly defined (nonholonomic). It is for these large systems that newly emergent concepts must be added to the physical concepts (without violating them) appropriate to small systems. The concepts must deal with the emergence of new form and function from the pieces that make up the system. (The emergence of new form and function in a dynamic physical sense is reminiscent of the mystic Hegelian-Marxian transformation of quantity into quality. It transcends that doctrine because it emerges from the deterministic nature of space, time, and physical law itself.)

It is the properties and 'laws' for such large systems that we must propose for biology.

6. Some General Concepts for Large Systems Analysis

It is of value to indicate major points of view which are the keystones of systems analysis and systems thinking, and the key to the discovery of systems relations.

(a) A system is not viewed as a unique singular entity (unless it can be the only one available to us - for example, the universe. Even then we visualize a potential collection of universes). Instead we attempt to visualize a potential collection of systems, which we would recognize as containing examples of the same system, i.e., members of the same set. These systems are not identical in temporal or spatial performance, but they possess similarities.

The reason that we require this identification is because metaphysically we cannot deal with the 'real' system. We have arrived at the issue raised by Korzybski. The map we draw is not the territory we draw from. The identity of the collection lies in our mind. The problem we face is similar to the act of knowing. How do we know a rock? It is the class of all objects that we identify with the properties of a rock. Thus it is the common possession of a few limited discernable properties that guarantees membership to this static class of names.

So it is for systems, it is the identification of a number of properties - some static, some dynamic - that is represented by our naming a system.

(b) Such an identification of a system practically carries with it the concept of the ergodic hypothesis of statistical mechanics (or more restrictively, the quasi-ergodic hypothesis). The quasi-ergodic hypothesis states that all regions of a hyperphase space[1] that are accessible to each visualization of a system are occupied with equal probability, so that the averages over the motion of one system in time are nearly equal to the averages of all systems over phase space.

This hypothesis is not provable. However, it is more nearly a tautology than a disputable thesis. What made the system 'knowable', was the ability to identify it in all of its aspects from time to time, independent of certain spatial and temporal transformations. This lumped the number of aspects that were the keys of identification into a fewer number. Another prototype of the system was 'knowable' because the same aspects were being carried along. One could, in principle, always embed the aspects of concern into a sequence of essentially equally probable recognized segments. Then the systems were identified as nearly 'identical', or possessing a common 'name' (namely, operating through a comparable concept). Thus the ergodic hypothesis is almost always essentially built into the identification of similar large systems.[2]

The systems may be a little different. Thus the ergodic hypothesis is a guide, not an absolute law. Its quality also depends on the size of the population. For example, there are significant population variations at the level of 200 million inhabitants of the U.S.A.; on the other hand, polls obtain 1-2% reliability in estimates at the 1000 sample size.

Our problem in discovering large systems relations is to embed the problem in a large enough universe that we do have a near ergodic hypothesis.

[1] The chains of a system are characterized by 'displacement' variables representing the degrees of freedom of the system. The hyperspace we refer to is made up of the displacements and their time derivatives.

[2] We may be puzzled by two teenagers, one who only studies, and one who only pursues members of the opposite sex. If we can see a substantial similarity of other functions - namely physiological - we consider their actions as ergodic. Of course, there will be differences.

(c) One way we may begin is with the view of a systems designer. He visualizes the kind of system he wishes to 'design' (e.g., a power plant). By design, we mean the selection of one or more major functions, 'purposes' for the system, and the development of a design theory for the system. 'Design' in the present systems context, will refer to design for a reasonable long sustained life system. Thus the design theory is concerned with its long term achievability.

To do this, it is generally necessary to visualize a series of homologous systems of different sizes that are capable of operation, i.e., the design is not tied to a specific size, but to a general principle that may be satisfactory to use to bracket the size of interest (e.g., the designer of megawatt prime movers is commonly limited by the number of principles for efficient prime movers he has available. He then has large scale constraints he has to meet. The biologist who came closest to expressing this point of view was D'Arcy Thompson.)

The systems designer may then require the development of all sorts of subsystems, which he may order up out of well-known arts (e.g., the components that are to be used for a large power plant are themselves rather 'canned' in their selection). What the designer is able to 'buy' are subsystems and mechanisms which fulfill his specifications without necessarily being tied to any one principle, or even sharply delineated characteristics. (If he wants temperature regulation, he can ask for it, not how it must be gotten. Thus if one animal uses skin sweat glands, another his tongue, another his tail, the master designer is not surprised.)

Disregarding all the intermediate details, it is generally only when most of the near static, or low frequency characteristics of the system are settled, i.e., what will assure that a system will persist, that the designer may review or have reviewed for him the high frequency characteristics, for stability, acceptability of performance, and sources of life degradability.

While all of this sounds remote from biosystems science, more like industrial engineering, it is the essence of all science applied to real systems. Imagine that you really are required to build a biological system from basic materials. How would you do it? The exact order of affairs may not be always pertinent. However, one may always start from the briefest epoch that represents its slow motion average. It is a period generally substantially shorter than the life of the system, but yet one that averages over the high frequency operative principle by which the system operates. It most often includes the primary slower regulators of the process. It is that period, in W. McCulloch's happy choice of words, for which "one day looks much like another," for it is these epochs that the ergodic unfolding takes place.

7. 'Directed' Chains - Characterizing a Large System

A large system has a 'purpose'. This remark is not meant to be merely a teleological statement, but one that will be given operational meaning.

By observing any system for a considerable length of time, we begin to infer, purely as hypothesis, some concept, some algorithm of what it does; i.e., what its one or more major 'purposes' might be. Such an algorithm, as hypothesis, attempts to cast light on what the future behavior of the system is likely to be.

Just as a large system has chains within it, the large system becomes a transfer apparatus or chain at a higher hierarchical level. As a result, it too may be degradative, etc. Our concern is with periodic or repetitive systems (if only one man had existed, it might have been fascinating for him, but perhaps difficult for nonexistent 'us' to have become excited). Thus, our study begins with systems which repeat, i.e., which have long life times with many repeated cycles, both in the appearance of the individual element and in its repeated forms. A biological system, like other large systems, is not an isolated entity, but a repetition of form and function in time. When we have achieved an adequate description of its form and function in space and time, then we can attempt to infer the purpose of the system from the complex pattern of behavior that we have observed.

In general, the description of a 'localized' or bounded system that exhibits repetitive behavior is not unique in a mathematical-logical sense. More than one classification or decomposition of its characteristics is possible. However, one decomposition that is always possible in an inconstant field is by means of harmonic functions. More precisely, any function that remains bounded in time (meaning that the excursions of the degrees of freedom of the system do not cover an infinitely large range), and whose excursions in any finite interval of time can be summed (namely, that there aren't too many wiggles to defy finite summation), can be uniquely decomposed into a sum of harmonic functions. If the function repeats indefinitely throughout all time, the decomposition may be expressed in terms of a harmonic (Fourier) series. If the function is not periodic, but defined for all time (for example, zero up to a certain time, then some aperiodic pulse or wave form over a limited period of time, then zero afterward), then the function can be represented by a harmonic (Fourier) integral. The former function is defined by a discrete line spectrum, whereas the latter function is defined by a continuous spectrum.

The descriptive adequacy of any experimental observations on a system arises when neither the form of the system nor data on the system are defined for all time. The uniqueness of decomposition comes into question.

The usual reason for seeking out decomposition by harmonic components or other well defined components is that if lines or bands can be found in a frequency spectrum which involve considerable energy or fluxes (as compared with a low level of energy or fluxes tied up over a very broad band - and referable to as 'noise'), then we often find associated physical mechanisms for these spectral regions. In contrast, transients may or may not be associated with significant physical mechanisms.

This decomposition property furnishes the reason and importance of harmonic analysis as the transition from the purely mathematical to the physical foundations for analysis of complex systems.

What kind of systems can we analyze? We can comprehend persistent systems, systems that show a repetition of form and function in time. The biological system qualifies for analysis by this test. What we have to do, physically, is to show that there are persistent harmonics which lock up energy or mass fluxes and we have a chance for physical analysis by 'spectroscopy'.[1] This path of analysis has described atomic and nuclear systems with fair success.

How do we know when a pattern is repetitive enough, and how do we show that we can deal with it over only a limited test period? Specifically, how do we know that a pattern will satisfy conditions for Fourier decomposition? We know only one way. We look at one system during a number of arbitrary serial intervals (not necessarily consecutive), or we use the quasi-ergodic hypothesis and look at n arbitrary systems over one such interval. We then perform the analysis that will locate its harmonics, if any. (Blackman and Tukey (3), a well known source, or Goodman (4) discuss the problems of analysis in distinguishing coherent 'signals' from noise.) However, recognizing that ultimately all systems degrade, this approach can be used only over segments short compared to the transient, aperiodic 'life' of a system.

Fortunately, the existence of a line spectrum can be established even with only limited, finite data segments from biological experiments. The value of harmonic decomposition has been enhanced by certain experiments showing strongly isolated line spectra in many physiological variables and processes. (Specifically, oscillations have been found in ventilation rate; in metabolic rate; temperature-thermal power oscillations; in blood gases - oxygen, CO_2; in blood constituents - red blood cell flow in capillaries, in blood sugar and blood lactate; in levels of some circulating hormones; in circadian rhythms of some biochemical constituents.)

Description of pulse-like events is much more uncertain. It is not that the events themselves are uncertain - they can be tested for in the same way, by the quasi-ergodic hypothesis, as can periodic phenomena, even though specification of the phenomena outside of the pulse domain is lacking. Nevertheless, spectral decomposition is of limited value for aperiodic phenomena. If understanding of an event is well founded, the event is preferably treated as an aperiodic physical phenomenon. The physical-chemical causality for the event is assessed, and viewed as being isolated. The fact that a long term genetic coding may have made the event possible does not contradict this

[1] By 'spectroscopy' we do not mean the more limited task of observing the spectra associated with atomic or molecular processes commonly via electromagnetic effects, but tne more general analysis of repetitive temporal processes in any system, whether 'atomistic' or macroscopic. This has become quite common in engineering, as part of dynamic systems analysis.

viewpoint. It is just that the boundary value problem of emergence starts in the time domain somewhat before the event, and one then attempts to seek out the more localized immediate causality.

8. A Statistical - Mechanical View of Life and Life's 'Purposes'

Generally, we can distinguish three stages in a system's life - its start-up phase, its developed, mature stage, and its degradation stage. To begin, we shall restrict our attention to the stage of mature development. During that stage, changes in an individual system are not explosive or one-time events. Observation and study suggests that a living system is an autonomous mobile factory that ingests energy and material, and conducts internal and external processes, so that it can continue to ingest energy and material and conduct these processes; it maintains its internal processes and form; and aperiodically it reproduces its own form so that its newborn can grow to a mature stage, where it in turn ingests and conducts its processes.

In overall design, a living system satisfies its needs in three time domains: a) within the high frequency range of metabolism; b) within a medium frequency range which is based on its 'experiences'; and c) within a longer-term frequency range for reproduction.

A statistical mechanical approach to systems starts, for its foundation, with a large ensemble of atomistic entities, some or all of which are capable of active, sustained transformations of energy. Because of force coupling between these entities they become self-organized, or externally organized, within a bounded domain. We now imagine a large collection of such similar ensembles. One such collection is made up of the system itself viewed at different times (e.g., the population of New York City in 1935 as compared with last Tuesday, or last Friday). We can proceed to compare all of these systems of the collection. One effective way is in a phase space made up of the distinguishable degrees of freedom of its atomistic entities, and of their rates of change in time. One such phase space locates the system as a point in that geometric hyperspace. Another such phase space permits that system, the point, to traverse the space as a trajectory. In that space, other systems in the collection can also exhibit their trajectories. We now propose to watch the collection of similar systems as they track that phase space.

Obviously the collection will appear as a 'cloud' of points distributed and moving through the phase space. For equilibrium states of the collection, the density of points are constant, or pulsing in time. For a feasible description of these collections, we require that the individual systems be not especially distinguishable; namely, that if we lose track of one system, and confuse it with another, then we can hardly tell the difference in the future course of the individual system. This requirement means that the systems behave with considerable independence of starting conditions, and with a considerable degree of repetitiveness of motions, which if not perfectly periodic are almost periodic.

If we restrict ourselves to lossy processes, we can postulate, with some justification from the work of Poincaré, that the motion of a system will appear as a nonlinear limit cycle in time, and as an orbit in phase space. At least a theoretical structure based on limit cycles, assuming well defined causal chains with considerable energy fluxes or mass fluxes associated with their spectral lines, might account for form and function in a system more successfully than other theories can do or have done.

A collection of systems in phase space operates with a variety of constraints, including the constants of their motion. Actually there is very little we can measure in the collection other than these constants, which we know as the summational invariants of the ensemble[1]. For example, in a physical system such as a nearly ideal gas in a container, there is very little to measure beside the constancy of its total energy and mass. As a result, its properties are assessed by application of statistical mechanics to a nearly constant energy distribution in phase space. The ergodic hypothesis (or various nearly equivalent statements such as that all accessible regions of phase space have equal a priori probability of occupancy for equal hypervolumes) may be invoked to characterize an equilibrium distribution of systems.[2]

Biologically, this statistical mechanical collection in phase space should probably be members of a species, or, perhaps only the adult members of the species. At a grosser level, with fewer degrees of freedom counted, the phase space may be a phylum; or with even greater loss in definition it may be the whole animal or plant kingdom; or it may be all living systems. However, at the kingdom level, the number of degrees of freedom are so drastically reduced that only some philosophic detail remains (e.g., all living things metabolize; their metabolism involves substrate oxidation; or animals metabolize three kinds of fuel, etc.).

The leap from the statistical mechanics of a container full of nearly ideal gas molecules to an ecological niche populated by human beings is a long one for the biologist to make. We should try to offer a route allowing smaller steps.

It turns out that the step from ideal gas to nonideal gases is already inordinately difficult (see for example (5) or Chapman and Cowling (6). The

[1] Summational invariants are those quantities in a system which are defined by laws of conservation during collisions between two or more components of the system. See, for example, Hirschfelder, Curtiss, Bird (5) for more detail.

[2] The statistical mechanical 'equilibrium' for a given system, illustratively a Maxwellian distribution of velocities for a simple molecular array, represents a sustained motion in which the particles or entities that make up the system neither decay to rest, nor go through a strictly periodic motion. The combined motion, however, is equivalent to a broad continuous spectrum of periodic trajectories. The individual system, a member of the collection of systems, thereby sweeps through the phase space with equal probability in each volume.

complexity is created by the internal degrees of freedom of the 'molecular' entities. We are confronted by a similar issue in the step from gas to living species. The species exhibit a still richer pattern of internal degrees of freedom. However, we find in any system to which statistical mechanics may be successfully applied that we cannot assign parameters sufficiently well to distinguish one system from another in an ensemble. We can merely identify 'summational invariants'. (Typically, in a simple mechanical system, there would be the invariants of mass, energy, linear momentum and angular momentum.) The many, many individuals or individual states that make up the system then are distributed stochastically in phase space by laws of distribution which involve only the summational invariants. We postulate that the very important statistical mechanical concept of the summational invariants in a collection of like systems provides a basis for defining the 'purpose' of systems, and, in particular, of biological systems.

If the summational invariants were known, then the equations of change[1] for a species ensemble could, in principle, be specified. These equations of change may then be used to describe steady-state[2] conditions for the species. When we imagine steady states for a species, we are imagining that the summational invariants are specified, and that the 'purpose' of the species lies in the maintenance of the steady-state. If the species is involved in a predator-prey oscillatory relation, the steady-state is the periodic type.

Consider now the 'purpose' of any individual human in a primitive culture. As an identifiable member of his species, he must move about in characteristic patterns in order to sustain his metabolic balance, i.e., in the long term he must conduct his internal and external systems so that his food intake and energy expenditure balance. If there are a finite number of physiological internal states which are freely accessible to change, then, just as in the statistical mechanics of mechanical systems, an equipartition of energy will take place in accordance with their weighted strengths. The system has to eat, to breathe, to void, etc. If the statistical mechanical idea is accepted, and if it stands behind the statistics ordinarily associated with observable physical behavior, as we believe, then 'energetics' in the human has to

[1] In the simple physical case equations of change refer to changes of total mass, momentum, energy, molecular species concentrations.

[2] In electrical engineering, two types of steady states are recognized: the steady state of rest, and a sustained state of periodic oscillations. One must note that the steady state of 'rest' for these statistical mechanical ensembles is an active equilibrium in which the underlying atomistic motions have achieved a stationary statistical character. The small gradient non-equilibrium thermodynamic theory then permits gross scale deviations from that equilibrium. Not too far removed from that equilibrium the system may thus exhibit steady state 'autonomous' oscillations; it may exhibit steady state driven oscillations; or it may exhibit maintained but displaced steady states, or transient state disturbances.

involve a partitioning of energy and switching of states among these tasks in a manner that will get them all done.

The descriptive problem that exists is that the switch mechanisms, whose actions contain the operative algorithms for the system, themselves partake of the energy partition. But it is easier to describe them as part of a communications chain inside the body, rather than the power chain. Nevertheless, even though small energy transfer may be involved in such biochemical chains, commonly because of their catalytic character, they require some suitable weighting in what may be described as the overall entropic processes of the body.

We may, in time, come to identify the biochemical chains associated with the macrostates of man. For example, we can already clearly recognize the stop-go states of motor activity in the human. The fight-flight-feed-copulate states have been long appreciated. With some confidence we can extend the list to include the fight-flight-feed-copulate-void-rest states. Beyond that, we can only ask, is there a 'necessity' for other modes of behavior? If so, can we find in their frequencies of occurrence the 'purposes' of the system?

9. On the Levels of Biosystem Science

The possible levels that could concern us in a biosystems science are:

(a) the chain of nuclear events - nuclear 'oscillators' (electrons, protons, neutrons, gamma rays. Note without these energetic entities, even now in the form of ionizing radiation, life would not likely have begun.)

(b) the chain of ionic, atomic, molecular events - atomic-molecular 'oscillators'

(c) meta-molecular macro-complexes - chemical oscillators - 'cells', organelles, and the subclass of C-O-H-N (plus Na, Ca, K, Fe, Cl, S, P) based oscillators

(d) a cellular community in an ecology, including organ systems

(e) the complex organism, including its subsystems

(f) a community of complex organisms in an ecology.

Biosystems science in a limited sense begins with the third level (molecular biology on the level of the cell and its subcomponents). However, we shall assume all science, up through the cellular, and concern ourselves only with three levels - the complex organ systems and functions in the complex organism, the complex organism itself, and its individual interaction within its surroundings and with its neighbors. The many other problems bordering these systems below and above in organization are at present only our marginal concern. The individual, we regard, as being ergodically embedded in his operative species.

Further, while there are three classes of problems that might concern us from a systems point of view, namely systems start-up, systems operation, and systems troubleshooting, in this phase we will only be concerned with one - a description of normal operation of the system.

10. The Normal Operating State of a Biosystem

We may start by noting the autonomous repetitive (i.e., harmonic) character of the behavior of an organism. An important characteristic that is not often made explicit about a living system, yet which is implicitly contained in this concept of repetition, is that the system is marginally unstable within its environment. This means that if the organism is put into a milieu, it will not remain still.[1] Conversely, if in a state of motion or agitation, the system will not continue so indefinitely when put into an isolated enclosure. In electrical engineering terms, we might say that its gain at zero frequency is indeterminate.

In systems analysis, we commonly start from an analysis of the D.C. characteristics, typically with an attempt to characterize the equivalent D.C. or - in extension - the low frequency network. It does not pose insuperable difficulty to recognize that the system may be indeterminate at zero frequency.

In order to arrive at a view of any lower frequency of a system, one averages or 'filters' out higher frequency phenomena.[2] For a system in which form and function are sharply fixed, the elementary scheme of harmonic (Fourier) analysis may be applied with not too much difficulty. It is much more problematic when the scheme is being used also to distinguish change in both form and function.

One attempts to average each remnant high frequency harmonic component (remembering that a system exhibits a repetition of function in time) to see whether the next lower frequency is sufficiently unchanging in form to regard it as a normal phase of the system. One may continue such averaging until the next lower frequency begins to change in form.

In rigidly formed systems, such as many nonliving systems, form doesn't change until one gets to the degradative state. Thus the low frequency response is often clearly separated from its changes in form (e.g., a home radio begins to degrade from the time it is first turned on. However, its electrical characteristics have not drifted seriously for a few years. Thus its low frequency response may be considered the last frequency whose period is shorter

[1] The possible metastable states of spores, or of chemically transformed virus are excluded from our present concern.

[2] The typical means is by using sensors which cannot respond that fast or with that detail.

than a few years. Typically this might be the warm-up time of a few hours. This illustrates how wide the gap may be separating the two time domains in some systems.) On the other hand, in a less rigid system, such as a living system (involving a high turnover of materials) the separation requires much more thought (e.g., a man may not differ much day to day - except perhaps near the days near death - but his structures begin to show considerable change in form over months).

In distributed systems there may be further difficulties. One must determine the states that the system operates in. We shall discuss these later on. These then have to be combined and subjected to boundary conditions. It is possible that nonlinear oscillatory states may emerge. Then steady state phenomena may or may not deterministically result from these oscillations.[1]

Finally, we begin to approach systems as complex as the living system. In a limited milieu, they may exhibit one pattern of excited states; in another milieu, other patterns. We find that the dynamic patterns of the internal systems become organized into overall behavioral patterns, perhaps best identified as organ constellations. To the degree that we can sharply distinguish the total exterior action set of the organism, we may refer to the external set (of states) as a modality.[2] Within this descriptive framework, what then is a 'normal' state description?

We are forced toward the following phenomenological position. (The same would really be true for a complex system known as a manufacturing plant, or a vehicular automobile, or a sheltering home as transfer apparatus.) We are concerned with patterns, considered as normal operational patterns, not patterns that result in rapid or near rapid death or degradation. Thus, we have to choose a phenomenologically feasible schedule of 'usage' that the 'manufacturer' recommends.

More specifically, one may arrive at the concept along the following path: Let the reader volunteer for a physiological experiment in which he is willing to commit his actions 'fully' to the experimenter's choice. What class of scheduling would he find acceptable for the following periods of time?

1 second
1 minute
8 hours
2 days
1 week
1 month
1 year
10 years
1 'lifetime'

[1] To illustrate with a living system, consider a reproductive cycle. Communications by a mating call or scent involves a risky distributed chain. Yet commonly a seasonal mating cycle will take place. One might expect a steady litter rate to emerge from this cycle, however other factors - nutritional, predatory - might upset this.

[2] By the term modality, we intend to capture the form or way of being, as apart from the substance, thus denoting a similar concept to that used in music.

At the high frequency end, physical capability itself likely governs. (You can't do thirty tasks in one second but you may be able to do five to ten.)

One should note, by giving the problem careful thought, that increased richness of state patterns are required, and lessening of rigidity of scheduling in order to make the longer periods tolerable. (Note for example the standard prescription for monogamous marriage that is commonly accepted, i.e., love, honor, obey, etc.)

It will be found that stable patterns of behavior - the modalities - can be characterized essentially as repetitive events that take into account internal oscillator systems - the organ constellations - plus repetitive cues that are supplied by the ambient environment. In addition there will be a certain amount of random events that can be characterized as 'noise'.

As an individual, are you willing to repeat a weekly schedule, regardless of how perfectly worked out, for 20 years without change? Most common answers would be no. Yet with a longer term break, such as 2 or 3 weeks 'vacation' per year - with its prescribed behavioral pattern - it becomes more common. Would one commit oneself for a two-week vacation every 49 weeks against a 52-week seasonal year? Such questions are harder to answer. Thus, a considerable scheduling by simple numbers of loose epicycles is the only a priori safe estimate. Of course, we can and do imprison man and beast and enforce more rigid choices.

Basically there is a quasi-ergodic character to the living system. Given a broad ranging ecology - even adding all of the modern technological landmarks that one is subjected to (this can be checked by noting one's own pattern development of habits in taking on a new life task in a new area) - regular patterns and habits develop. However, the behavioral constellations that any one person arrives at are similar to those for most other people.

Thus, the normal operating state of a biosystem must be defined within a representative schedule of motor performances programmed in time to resemble in ergodic richness and scope the performances of most other examples of biosystems in that ecology[1]. The ecology itself may not be an unchanging environment (else the reality of sensory deprived states and studies will become evident), but a vicissitudinous changing impulse spectrum of inputs. Light and dark, and rain, and hills, and water and earth, pestilence, warmth, cold, meeting friend and foe, life with all its comforts and terrors are the common experience.

[1] While it is difficult to put down a sequence of random numbers, it is possible to artifically choose numbers that at least resemble a random number sequence in many significant ways. The same is true for a choice of a 'representative' physiological schedule.

One must note that external motional responsiveness - its autonomous motor character - is a fundamental part of the characterization of living systems. Of course, the biologist should not be immediately misled by the tree, or sea plant standing idly and letting the world come to it. This overlooks simply the time and scale of motion. (It is really only the indeterminate dormancy of spores which requires review.)

11. Defining the Normal Response of a Biosystem

The normal vicissitudes of the milieu are not so rapidly changing or so violent that the biosystem is wildly driven. If wild perturbations happen too often or too violently, the system commonly breaks down. (Namely, the system is essentially designed with a limiting frequency.) Thus, more commonly, system behavior unfolds against a slowly changing vicissitudinously impulsive background. (The graininess is considerable, but of small amplitude. The hills and dales and fields that are encountered are gentle and rolling.) It is against this background that we find the living system to be particularly viable. (Else it is quite hung up and limited to borders, or only sparsely distributed.[1])

Normally, we find a preservation of form and function, and of conditions in the interior milieu in spite of a turnover of materials. The dynamic regulation of this state by mediation of many limit cycle oscillator systems has been renamed homeokinesis (4), as a preferred modification of homeostasis, the regulation of the internal environment.

Thus, to a considerable extent, our task will be to place and question the many process and mechanism chains by which homeokinesis emerges, as both temporal and causal sequences. In toto we hope to represent the many components of the normal response of the system. Each regulation, when it emerges, is most generally associated with a spatial and temporal scale.

It is useful to provide the reader with a unitary scale to think about, along which he can imagine the patterns of behavior both internal and external emerging melodically as the winds of the milieu play upon the organism. We thus offer him, to start with, the following introductory but undocumented representation to help fix his attention on the phenomena that will interest us.

[1] These statements do not mean that all of life is serene. The systems are 'designed' to be capable of handling particular tumultuous events. Witness the spawning trek of the migrating salmon. However, in general, the average metabolic level is far below the peak metabolic level. Thus violent activities tend to be somewhat infrequent.

Time Scales

(either aperiodic impulses, or periods)

1-5 millisec.	Short range nerve impulses
30-100 millisec.	Long chain nerve or coordinated muscle impulses
0.3-1 sec.	Coordinated CN complex
5-15 sec.	CNS neuro-endocrine commitment to a system 'posture'
30-100 sec.	High speed endocrine pathways
300-500 sec.	Common fast endocrines
20-90 min.	Whole body cooperation for a major task segment (e.g., listen to a lecture)
3-5 hours	Whole body commitment to a modal complex (work all morning)
24 hours	Recharge cycle for the whole system - its primary relaxation oscillation (generally cued by the day)
3-5 days	Follower endocrines for regulation of set points
15-60 days	High frequency domain for life postures and for longer term endocrines that have to perform specialized life functions. Adaptive processes of the whole body.
8-15 months	A period that is largely cued by the year and its seasons. Most often long range endocrine actions lead to periodic processes that are genetically locked in.
6-20 years	Life programs. It is the time for mastery of a whole life complex, including a large number of skills.
50-80 years	The life period, between initial growth and final degradation.

Degrees of Freedom

(We present here four different hierarchical levels of description, and the associated entities that can be found in different measure states. Somehow we expect that the degrees of freedom of the biological organism are to be associated with these measures.)[1]

1. Chemical Materials

   Metabolic: H_2O, fuels-fat, carbohydrates, proteins; O_2; CO_2, lactate.

   Metabolic by-products and intermediate products: CO_2, lactate.

   Cations: Na, K, Mg, P, Ca, Fe.

   Anions: Cl, HCO_3, PO_4

   Immunological reactions: Not listed (too many to count).

   Hormones: Not more than a dozen major ones.

   Enzymes: 2×10^5 species.

2. Enabling Structures

   Cardiovascular, microvascular systems, liver, kidney, brain, spleen, nervous system, muscles, senses, endocrine glands.

3. Processes

   Electrical excitation and inhibition, materials exchange, fluid transport, chemical signaling, materials storage, materials manufacture, information processing, power transforming (motor systems).

4. Behavioral Modalities

   Ingestion, sleeping, working, sex arousal, anxiety, anger, fear, interpersonal involvement, internal involvement.

[1] Physically, the degrees of freedom are those generalized displacements that are found to contribute to changes of energy of the system; more generally to the summational invariants of the system. They include geometric displacements, electrical displacements (charge), and chemical displacements (concentrations).

12. The Phases of a System - Start Up, Life, Degradation

In somewhat more detail, we must understand what a systems description must amount to. We assume that the individual system has many degrees of freedom, and that the system states respond to the changing internal and external milieu (since they are made up of linearly unstable but nonlinearly stable oscillators, this is no surprise). Thus a rich temporal patterning arises. The class of all possible displacements and their rates of change is rather richly filled, and as we note it in time, we find that the rather rich dense filling only fluctuates moderately. Our problem is to find a decomposition for the description of these state changes.

First we identify three phases in the system - start up; normal 'life'; degradation. Experience with many kinds of systems suggests this as a useful initial decomposition.

Having chosen a quasi-equilibrium systems physiology for a first round, we limit ourselves to the 'life' phase. This for humans is clearly post-puberty (say 15 years of age) to at least the age of female menopause (say 55-60-65 years of age).

Within the life phase, the active modalities of the system are limited in number and their range limited. Thus in any individual there is considerable cycling among states. This gross space filling of a limited number of states is essentially sufficient for a well-bounded and well-defined average. A system - human being or otherwise - cannot be running at high speed all the time, nor can it be resting all the time. This, plus the bounded ranges, leads to well defined averages.

The ergodic or quasi-ergodic assumption essentially assures that a well filled unimodal distribution of states around the mean will be found. Thus from a fairly large population, it is not surprising to be able to select a stationary distribution that will account for the states traversed by a large number of the population. The action of many individuals observed over many intervals of time is like that of many other individuals, and thus like that of themselves at various times.

Thus we can visualize a common range of expected behavioral range and performance of 'normal' physiological systems. While this is stationary, it is a stationary distribution, but not of fixed states. Thus we may visualize a large number of schedules which have essentially equal stationary measures - of say means and ranges. We can differentially detect other measures - such as higher moments - but these involve more information than a system warrants for a first round. Thus we are satisfied with a phase space that brackets the mean state and the state range (for both displacements and rates of displacement).

It is for this 'normal' physiological space that we are required to identify the systems relations.

We proposed that for the behavior of an ensemble of individual systems, we can note the following behavior:

(a) A time averaged stationary behavior.

(b) Behavior driven by frequency cues. (For biological species, this involves generally at least the geophysical day and the year.)

(c) Behavior representing autonomous oscillations.

(d) Behavior representing 'random' stochastic variations.

It would be our thesis that the individual dynamic behavior consists of the (a) plus (b) plus various branched (Markov) chain processes involving (c).[1] Basically, it would be our contention that for the individual all selections of branches from (d) which tend to create richer motional patterns in phase space without particularly changing the zeroth and first moments for the population measures in phase space are feasible. Thus occasional deviations from normal stationary population measures are possible for the individual (a person will likely get sick, etc.). However, the very definition of normality will also imply sufficient long life for the individual to persist for the expected life span.

Our first step in systems analysis is to determine, within a hierarchical level, what relations govern overall behavior on the average and then over a time scheduled average behavior. This has to be consistent with the likelihoods of (b) and (c), and in principle is a behavior that can be repeatedly obtained in a test pattern.

13. The Point of View of Systems Design

Let us examine what we have been saying within the scope of systems engineering. First take the point of view of the systems designer. He knows he must get a system or subsystem of some larger system to perform a certain high frequency function, and he must perform that function for a specified 'life' of meaning to the larger system. (It is not necessarily the entire life of the larger system.)

This specification generally leaves him with four kinds of problems.

[1] A typical daily chain for human behavior might be sleep-wake-void-eat-work-eat-work-eat-rest-play-sleep. With some moderate variations, to a large extent this represents a sedentary man's behavior in a technological society.

(a) He must choose the operative principle and 'design' the major component that will perform the high frequency function.

(b) He must get a large host of ancillary components designed, or procured 'off the shelf'.

(c) He must assure that the components can be brought together so as to assemble the overall system.

(d) He must assure that the components and the overall system can meet life requirements.

The design and reliability of the central component is a major problem. Thus for the cardiovascular system, the higher frequency beat by beat operation of the heart, and the lower frequency carotid sinus baroreceptor regulator are undoubtedly the high frequency components that performs the major function of providing a high pressure source of oxygenated blood.

However, equally severe a problem is meeting the long term requirements of his system. It is here, generally that the systems designer spends his major time. The other problems, the major component, and the ancillaries, he leaves to assistants, or for second choice.

It is this very difficult, more integrative, lower frequency view we are attempting to have biologists accept in a biosystems science. We start from the view that their training has given them an idea of much of the details of the high frequency response element. Now we want to broaden the outlook to the overall systems dynamics.

II. THE ORGANIZING PRINCIPLE OF COMPLEX LIVING SYSTEMS

In the first section, we provided some background to physical thinking about systems, and a preliminary clue to how we might think about the biosystem in those terms. Now in this section we intend to provide a more coherent introduction to the overall design of the biosystem, of both its physiological and behavioral mechanisms.

1. Instability and the Nature of Life

As a suitable engineering definition, we may provisionally define a life-like system as any compact system containing an order and distribution of sustaining nonlinear limit cycle oscillators, and a related system of algorithmic guide mechanisms, that is capable of regulating its interior conditions for a considerable range of ambient environmental conditions so as to permit its own satisfactory preservative operation; that is capable of performing these preservative functions for a long period of time commensurate with the 'life' of its mechanical-physical-chemical elements; and that is capable of recreating its own internal systems, or being recreated, out of materials and equipment at hand in the ecological milieu.[1]

An essential characteristic of a living system is its marginal instability. It is this characteristic which leads to the biosystem's oscillatory nature. Its principal dynamic properties are that it hungers, feeds, and can move or creep about so that it can continue to hunger, feed, and move or creep. At the right unfolding time, it couples and reproduces so that the newly formed unit can hunger, feed, and move about. Both the external and internal environment constantly present the organism with an impulsive (vicissitudinous) input against a background of the slowly searched, changing milieu. As a result, the motor systems of the organism are plunged into intermittent search modes to satisfy all of its hungers.

[1] Similar to the character that we have identified with the entire complex organism, we incline to agree with Waddington that it is unlikely that anyone before Brian Goodwin conceived of or argued so forcefully that the biological cell is an oscillator system (7). Nor should the contributions of Strumwasser be overlooked in highlighting autonomous biochemical oscillations in nervous systems (8).

It is pertinent to recognize the most primitive function of nervous system elements. The primary function of neurons is to make connections. They do this by making protein. Phylogenetically, in simple forms of life, the primary activity of cells that act as controllers of other cells is secretion. As the cellular processes grow longer, one then begins to see the occurrence of high speed signaling. However, these cells basically never lose their action of secretion. It is most apt to note this "chemical materials for export", Strumwasser's elegant phrase, doctrine receiving new impetus in a modern form through the discovery of long term temporally periodic processes emerging from the nervous system.

Finally at the lowest level, recent significant contributions, from Zhabotinsky and Katchalsky, have demonstrated the possibility of chemical oscillators. Thus the laying of foundations for a dynamic theory of life processes is well under way.

Internally, there are many active biochemical chains organized around the emanations from cells. These chains are parametrically unstable, and the changing internal milieu provides ever-present input excitation of their state. We use the term chain to stress that we are discussing causal chains of events and elements that are not absolutely fixed spatially. Many of these chains in fact have external links.

Functionally, at all levels, the key principle by which the living system is organized is dynamic regulation of its internal degrees of freedom (concentrations, potentials, fluxes). This is achieved by the mediation, mainly by inhibition or release from inhibition, of a manifold of oscillatory (or rhythmic) processes which make up the many biochemical chains in the organism. It is by a number of modulating schemes - generally asymmetric - that these processes change and regulate the mean parametric states. These regulating chains tend to be organized hierarchically at a sequence of temporal levels.

Structurally, what is common among the internal organs is organization of their actions into essentially closed chains of biochemical-mechanical-electrical nature, involving the solids, liquids, and gases in the body; e.g., the breathing, heart beat, voiding chain. The systems exhibit complexes of stable, rhythmic limit cycles, often passing through transitory stages as the organism is affected by changing contingencies in the external milieu.

The function of the central nervous system with its memory, communications, computational, and learning capabilities, is to provide algorithmic content capable of mediating the stability of the internal chains, so that a satisfactory pattern of behavior emerges among the organism's many behavioral modalities (such as eating, voiding, sleeping), perhaps twenty in all in man. It modulates the systems into modes.

For the scheme of regulation by which these oscillator systems are modulated through their nonlinear stable operating range we have proposed the name homeokinesis. It is chosen to stress and extend the principle of homeostasis by which regulation of parameters in the internal watery milieu, independent of the vicissitudes of the external milieu, is the condition of life. In extension, it is only by the manipulation of kinetic variables of space and time - namely, by mediation of the limit cycles - that the scheme can be achieved.

What results is a behavioral patterning whose richness depends on the topology of the central nervous system rather than on its size, in which the successful organism threads its hungers (or needs) so that the emergent patterns fit the ecological-ethological environment and the modalities of the system. Behavior is essentially periodic, episodic, and repetitious. Individual or species failure may be marked by patterning that saturates, or that goes into violent oscillations with regard to its hungers.

The motor-actuated living system unfolds its states, moment by moment. In each posture (the action of the body on the body), the system is temporarily locked into an orbital constellation of all of its oscillators. The psychological-physiological 'moment' then changes from instant to instant. The biological spectrum emerges from the many oscillatory chains that enter into the constellation.

Salient characteristics of the system include, first, the turnover of material. Although form and function tend to be more fixed, the actual chemical constituents of the biological 'factory' are all in flux. Second, there is a fixity of form and function which can be identified at three levels: (a) the metabolic high-frequency domain in which the operating chemical chains exhibit their character as modulations of a mean state; (b) the epigenetic medium-frequency domain in which the fixed genetic coding unfolds with links that can only be formed from the content and experience derived from the external ecological milieu; and (c) the genetic frequency domain, which is rate governed with a long-time relaxation phase and a short escapement phase, in which chemical coding exists for the dynamic gating of catalysis by which structural-functional hereditary chains of reliable reproduction emerge.

2. The Spectroscopy of Man - Organization in Space and Time

Study of such a mobile dynamic factory as the complex living system may be tackled effectively by dynamic systems analysis. This requires a technical decision as to what processes are of concern at the shortest time and the smallest spatial element of interest, and at the longest time and the largest spatial element of interest. Within these limits, spectroscopic analysis of spatial and temporal effects can be attempted. The lower functional levels are illustrated by the distributed capillary bed in which near 10-micron free red cells interact with 1-micron capillary system walls; by the glomerulus; and by the distributed communications unit in the local neural net. The lower time scale limit is about 0.03 sec. Extensive electroencephalography of the brain cannot find significant organized content below this limit.

At the upper limit, the isolated animal is our unit. Beyond lies the province of the social sciences, although behavior begins with the unit-to-unit interaction. The long-time limit is the single 'relaxation' process that is a man's lifetime. More restrictively, it is the failure of one or more internal systems that marks the beginning of a degradation phase. The temporal spectrum of behavior is centered, grossly, much closer to 30 days than to a lifetime; yet the 15 years from birth to adolescence must also lie within focus. However, it is necessary only to note the existence of the family unit of mother-father-child and of the peer group of the individual.

In addition to sustained spectral elements that exist at the gross organizational level of the organism there are a number of aperiodic flaring instabilities that govern the individual's history. These are determined by maturational cues to the central nervous system. The first such stage is that of birth and living within the mother-child constellation. The mother teaches the child routines of living by whose repetitive patterning the child can survive. A plastic brain develops and encompasses a multidimensional image of the body - its interior and surface - and the external world. Guide rules emerge which lead to successful patterning of the rhythmic modes that are represented in the body image.

In a subsequent peer or chum stage, the youthful organisms play and practice routines until they are well adapted to the milieu. Interpersonal forces begin to develop the socializing constellations that are a condition for sustaining life. At the right time of epigenetically emergent adolescence, male-female constellations are formed that lead to reproduction.

At the macroscopic level, why individuals are bound into orbital constellations that make up the content of bio-,psycho-, or socio-spectroscopy is a mystery. A model is suggestively found in the quantum mechanical system known as exchange forces. An individual projects his body image - totally or in part - into the other object's shell. This body image of self and the physical image of the other are internally compared, for their complementarity or their congruent character. From this exchange of body image arises empathy, antipathy, indifference, all the shadings possible of interpersonal force. It is a binding to be likened to the analogous quantum process that binds atom to atom as a shared electron cloud.

At the atomistic (high frequency, small spatial field) level of the biological system, the level of the cell and its emanations, the essential processes are the genetic coding and replication, the energy-releasing chemical reactions, and the complex of catalytic reactions, essentially enzymatic, that lead to synthesis of materials for both structural form and function. The unit processes involved are, typically, transport, conduction, convection, diffusion, bi-polar stress, and chemical linking. However, the mechanisms of oscillating biochemical reactions (9) have not yet been fully determined.[1]

The spectrum of the actual biochemical chains that exist in the biosystem is not continuous, nor even densely populated. Instead, it appears that there is a rather limited time fracturing (or time locking) around which processes tend to form and be cooperatively involved. There apparently exists a rather limited finite matrix of regulated elements. Its columns are elements given by the metabolic reaction (fuel, oxygen, water, carbon dioxide), some other chemical constituent streams, typically electrolytes; its rows are time scales. There seem to be regulating chains that fit many of the temporal intersections. Not all animal species use exactly the same chains or time scales, but the density is similar.[2]

Within the space and time domain that represents the individual, his characteristics can be presented as a spectrum, page 29. It is the analog of spectroscopy for atomic or molecular structure. (The schematic figures grossly specify the essential regimes.)

[1] It is only the subject of oscillating chemical reactions that is off to a start at this time of writing.

[2] While this covers the functional portion of the matrix, it neglects the structural portion of the matrix, in which many steps of synthesis take place by which structural form and functional specificity are assured.

3. The High-Frequency Biochemical and Electrical Foundations

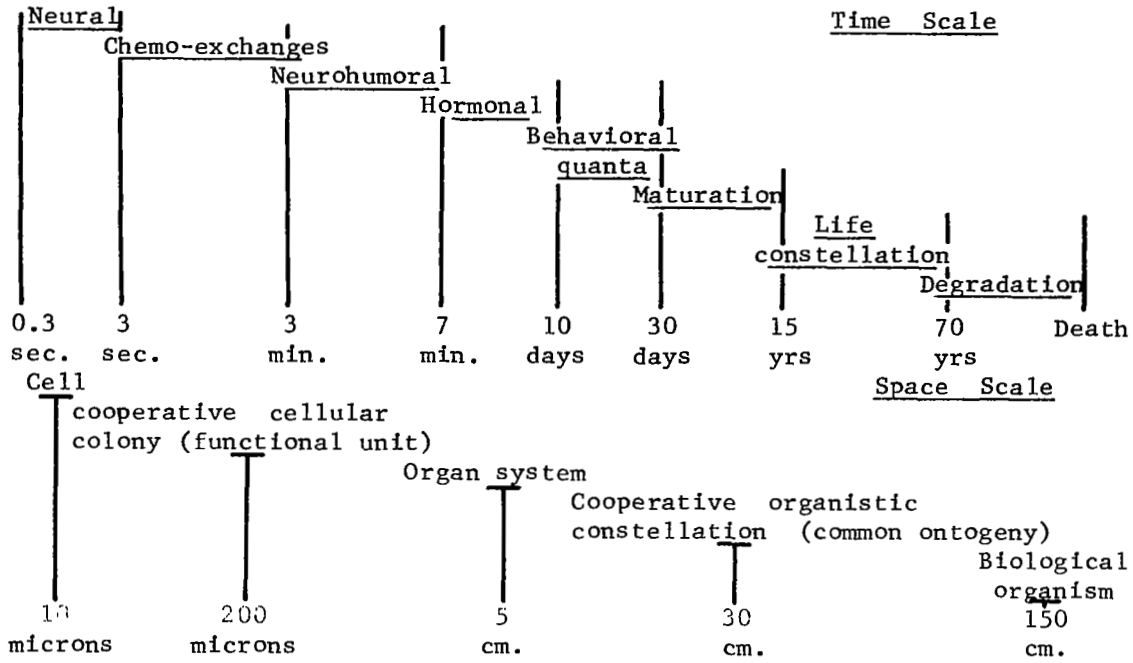
An introduction to regulation and control in the complex biological system may be usefully viewed in terms of a series of polar concepts. The hierarchical control systems are chemical and electrical. In the main, control is exercised through an electrically sensitive structure - the membrane. Roughly speaking, chemo-electrical mediation of the membrane controls electro-chemical processes within the membrane. Within the complex biosystem, the two major coordinating membrane systems are: the capillary membranes, controlling chemical flux exchange across the capillary walls; and the nerve membrane, controlling 'informational' flux exchange (of electro-chemical nature). Within each specialized cell, there are of course many other membrane structures.

Communication in the organism goes by two channels: 'to whom it may concern' messages go by hormones in the blood - i.e., by chemical signals; and specific messages to particular places go by nervous impulses - i.e., by electrical signals. The latter, so far as housekeeping controls are concerned, are mediated by the antagonist autonomic systems, both of which use acetylcholine in intermediate relays; but only one, the parasympathetic system, like the motor nerve, uses it in the target organ. The other, the sympathetic system, uses it peripherally in two ways: one locally, in the target organ, and the other by excreting it from the adrenals into the blood stream, like a hormone - which has a circulation delay of some seconds before it reaches its target and, typically, a release time of several minutes. Together, these half-systems control vasoconstriction and dilation. Their principal central controls reside in the hypothalamic and deep centro-encephalic structures.

As a rudimentary description at this level, signals received from the milieu - both internal and external - excite the nervous system into transient states. Its chemo-electrical communications signals modify the stability of chemical chains via the endocrines or neuroendocrines. Follower characteristics to the nervous system state emerge to support the motor excitation that provides the system's movement. In longer term, it is actually likely the persistence of chemical signaling - starting from conception - that forms the algorithmic content of the nervous system's transient response which is identified as behavior.

Without the nervous system mediating the biochemical chains, their predominantly unstable characteristic would show up. The nervous system is thus the mediator between information - a measure of contingencies in the surrounding milieu, both internal and external - received as sensory input, and the unstable motor and glandular systems. The resulting pattern is a controlled synchronization in which most systems are regulated in a limited oscillatory range by repression or inhibition, while some are released from inhibition and spring into orbit. System motion proceeds by the scheduling of these orbits in time, unfolding its repertoire of behavior moment by moment.

While higher frequency rhythms - including brain 'waves' (EEG) of the cortex and microvibration at the body's surface, as well as the aperiodic pulses and trains of pulse-like potentials associated with nervous paths - are sometimes overlooked in considering the large-scale and long-time integrative



Dynamic Spectrum for Man

action of the system, they are always present. They involve chemo-electrical escapements (8). They utilize mechanical and hydraulic links. Their presence makes it impossible to view the system, at any level, as static. However, conversely, they are not the communications language of the system. Their buzz is not the information flux, only an indication that the system is dynamically in flux. Decoding the internal communications and transport fluxes is a difficult matter.

The informational influx arrives from about 10^8 transducers for electromagnetic waves and about 10^6 for chemical and mechanical reception. Some of these have short delays before initiating millisecond pulses in the first input neuron, but it is trains of about 7 or more impulses in each of several parallel neurons that constitute its signals to subsequent neurons. Hence the spectrum begins at about 100 Hz. The fastest closed loop through the muscle and brain takes twice as long. Cortical tasks, like locating a spot of light, require a moment of about 0.1 sec. Similar informational fluxes from internal transducers course through the nervous system so that some of the subsystems of the brain, notably the hypothalamus, are informed by these signals from the heart, lungs, gut, arteries, both baro- and chemoreceptors, as well as the body's muscles and joints. These informational winds of the milieu, external and internal, coursing through the brain, update its model of the body and the external world at approximately 10 Hz. Posture and movements induced by muscular activity are reported by the electrical coupling from mechanical, hydraulic, and perhaps electrical or chemical receptors. Information is yielded in two ways: tremors furnish the information as to posture; and, in movement, the information is derived from the motion apparently accompanied by an inhibition of the tremor.

The hypothalamus is the highest automatic regulator of the potential state variables in internal organs, glands, and the blood vessels. It patterns the response of autonomic systems at a medium-cycle time scale of minutes. This includes hormonal control through the pituitary portal system, forming a switchboard that operates in a ring oscillator mode, cycling through its affector variables. It acts as a slow follower on chemical signals produced within those various internal systems.

A most significant signal is the socially conditioned high-speed epinephrine signal, generally seen as the anger, fear, fight, flight autonomic response from the adrenals. Another component of this adrenergic response of the nervous system, high-speed release of norepinephrine at nerve ends, outlines the system at the one-tenth second level. This is used as an 'arousing' signal or 'groping' signal for what the instantaneous properties or status of the system is, and what may likely be the fact of the external world. As a sustained small-signal excitation, it helps keep the internal systems sufficiently regulated to face the motor system demands that may be placed on it by the command system, and to ready the system for follower action.

It is the reticular core furnishing discrete informative signals to the rest of the nervous system - in the 0.1 to 0.3-sec. range - that provides the sustained groping signal. The adrenals furnish large-scale regularizing or follower signals throughout the blood system at the 1-2 min. level. The hypothalamus then provides follower action at the 5-10 min. level.

It is generally accepted that the reticular formation likely is the large-scale arousal system in the brain. As a parallel core system throughout the neuraxis and midbrain, it has connections to all levels of the nervous system. Acting as a probabilistic computer based on all internal and sensory inputs, it controls the arousal of the system with regard to the direction and level of attention and of motor activity.

The most meaningful hypothesis is that signals from all sources run up and down the reticular core which contains the potential command system to commit the organism as a whole.

High-frequency behavior has characteristics that can be readily seen by watching any animal. The system is basically unstable. If you put the organism down, it will soon start to move. If it exhibits extensive motion, in time it will stop. Whatever the inputs, whether nearly constant or changing, the system changes. It changes its postural dynamics. Furthermore, men do this routinely at a rate of approximately 10 per minute. These changes can occur within a few nervous 'beats' - "moments" of up to 0.3 sec. Once developed, whether by learning or otherwise, these responses are subcortical. (In the response of the body to the external world, they may retain and require a cortical contribution, as illustrated in speech.)

Illustrative of postural elements are: postural attitudes of parts (head, feet, body, hands); dynamics of parts (e.g., scanning any scene for movement); characteristic movements (yawn, stretch, rapid eye movement, laugh); insecurity stereotypes (tics, twitches, scratches); sweat response; saliva response; voice pitch.

At high speed, the reticular system can act on displacement, velocity, or acceleration inputs within 0.1, 0.2, or 0.3 sec. Its core can obtain information and provide the command for about 15-20 kinds of actions. Logically, the reticular core operates as an abductive system (in the Aristotelian sense). It commits the system. It questions whether an information state is a case under one rule or another and takes a decisive action that actuates the controllers of controllers; i.e., it is a command system. In doing so, it may or may not consult the cortex.

The subcortical dynamic postural elements make up a manifold of responses. A classification of the chemical foundations of those response elements, for example, among the endocrine systems involved, does not exist as yet. In a subsequent section, we will assay a very crude description.

The subcortex sets various internal systems into orbital paths. The three large slower follower systems seem to be:

(a) The adrenalin-blood system. A major follower element seems to be the oxygen flow available to the tissue (as marked by the red blood cells) through the capillaries. The flux wave, involving concomitantly other metabolic elements - sugar, carbon dioxide, water, lactate, heat production - in dynamic cycles, likely represents the 'escapement' for the thermodynamic power cycle of the system.[1]

(b) The hypothalamus (at the minutes level, with a connection to the pituitary system as the master regulatory endocrine gland).

(c) The pituitary gland. Actually one may regard the hypothalamus - pituitary, pancreas, and adrenals as a systems constellation that acts in this slower time scale to target on the motor system.

In the lower animals, the reticular system contains the entire executive logic of arousal and shutdown of function. In the higher animal, the reticular core assigns the (phylogenetically) newly emergent role of induction to the cortex. If the input pattern is not a case that is fitted by standard analog patterns immediately available from the basal ganglia, the case is referred quickly (e.g., within 0.1 sec.) to the cortex. The cortical memory (whose storage place is as yet unknown) is an analog memory of many past cases. Presented now with external patterned 'facts', it 'guesses' at a law (i.e., it wires together a network response that excites the motor oscillators into action). The cortex 'takes' habits. Once these are set up on the motor side, the cortex is often

[1] The physical view of organized energy transfer is the task of thermodynamics. Such transfer involves a sustained temporal process and, thus, an escapement, unless tied uniquely to cues such as light-dark, or seasonal changes. Simpler biological systems that are tied only to cues may exist. However, circadian rhythms have been demonstrated in very simple biological structures. The complex human, homeokinetically operating, is freed from the vicissitudes of the milieu. Thus it must be self-timed.

no longer involved and the solution analog may be transferred or formed within the basal ganglia.

In primates, the cortex is very busy with optical signals. This is the 'price' the biological system pays for a precise invariant field. (Other fields are not outlined with such detailed precision and consistency.) In man, it is also quite busy with verbal signals. There are well defined regions associated with speech.

To attempt to adequately describe the central nervous and endocrine system components, and the currently known or suspected functions of the divisions of the brain and endocrines, is outside the task of this introductory section.

4. The Medium-Frequency Range of Behavior

We now approach the medium time of behavior - the 1 to 1000-hr. time domain. This is the time to go on vacation, get drunk, fall in love, take a job, get an important idea, get married, menstruate, commit suicide. Its unit of cueing will have a time constant of the order of the female menstrual period.

Common in philosophic-psychological speculations are aspects of the mind-body problem, such as the distinction between responses that may be closely coupled and 'wired' from input to output by an interposed, mechanistic system, as compared to a more loosely coupled indeterministic system. (See for example Cobb (10)). Much of the argument can be avoided at this point if it is recognized that our present concern is the result of changes in the state of the internal inputs to the central nervous system, rather than to external inputs. In the time scale of concern at this point in the discussion one day is much like another, and one week like another.

Behavior is drawn as patterns, it seems, from the concurrent connected internal states of the biological system given in Table 1. They may be identified as dynamic action modes of the system, such as "The system sleeps."

An essential characteristic in this longer time domain is the instability of the nervous system. The animal tends to develop an internal rhythm dominated mainly by endocrine chains arising from changing threshold cues; the actual performance is conducted largely at a subcortical level. Social activity and intellectual activity all are ritualistic past the minor originating cues.

A most significant element in the organization of behavior is the use of cues. (What should be noted is the large number of geographic and social cues.) In case of a temporally or spatially cued input, the animal will develop a ritualized behavior. (By definition, a ritual is a patterned response which

Table 1

Action modes	Percent of time
Sleeps	30
Works (motor activity directed toward ends)	25
Interpersonally attends (body, verbal or sensory contact)	8
Eats	5
Talks	5
Attends (indifferent motor activity, involved sensory activity)	4
Motor practices (runs, walks, plays, etc.)	4
Sexually involved	3
Rests (no motor activity, indifferent internal sensory flux)	3
Is anxious	2
Is euphoric	2
Drinks	1
Voids	1
Angers	1
Escapes (negligible motor and sensory input)	1
Laughs	1
Is aggressive	1
Fears, fights, flights	1
Envies	1
Greeds	1
Total	100% ± 20% of time involvement

is stereotyped.[1]) It likely follows from the system instability. An animal cannot maintain undischarged nervous excitation energy, but must seek to release it by releasing the inhibition on some motor response system so as to unstabilize it into orbital action.

What is particularly noteworthy in the biological system is the patterning of behavior that runs it through a repertoire of performances. Psychology of higher animals should be regarded as a 'pattern' psychology. Since the patterns are not rigidly fixed, they are not deterministically preprogrammed and not rigidly cued; they can only be self-actuated, and thus must arise from internal instability, so that the system goes into 'motor-sensory-internal organ' motion. However, the orbital synchronous patterns that arise must thread, in an ergodic sense, all of the 'needed' systems responses. These 'needed' responses must be regarded as 'hungers', metabolic for individual survival, and genetic for species survival. This temporal threading must fit the cues, or the cues must have been fashioned or adjusted or adapted to fit the species. Then it can exist.

A species must be 'comfortable' in its pattern fitting into the time space. The first nonlinear role of behavior (likely for higher cortical species) is that a noncortical routine of patterns must be achieved that fits the cue space with small integral cycle numbers. Then processes - the hungers, particularly - *come off on time and the system is not in sustained stress.*

For such compatibility to be arrived at, there must be a sustained effort, by orbits, to bring the existing set of body images toward some optimum. It is some optimal view of reality that the mother and father or their surrogates attempt to develop in the child. It is a view of "Which way is up"! (11).

What is really basic to the plastic higher brain is the alternation in state between two essential roles of inward behavior; for want of a better name, an anxious or dysphoric state and the euphoric state. In the higher animal, it is the cooperative impact of all signaling interfaces that produces the set for the longer time scales. The periodicities may be in the few per day to one per 2 to 3-week time scale.

5. The Low-Frequency Range - Life's Foci

At even longer time scale, the predominant note is the great flaring maturational instabilities. The first is that of birth itself. In the first life phase, the newborn infant faces the development of a routine of rhythms whereby its greatest sensory interface - the oral interface - is encompassed within a 'satisfying' schedule. This develops in the brain as a dual mother (or surrogate) - child 'symbiotic' oscillator system, a constellation. Anxiety

[1] The stereotypes of behavior may be taught by the mother. In rough approximation, they recapitulate the phylogenetic progress up to the particular species (Gesell).

follows upon euphoria; motor patterns stuffed with ritual fill develop. The unfolding maturing nervous system brings new sensory interfaces (anal, gastro-intestinal and urethral, genital, motor, visual, kinesthetic, etc.) into the field. The plastic brain encompasses these. Priorities and patterns emerge.

The second lesser flare occurs when the system 'masters' its primitive routines. There is spare computational capacity. As an integrated "Gestalt", the system grasps its freedom from the immediate 'mother-child' milieu. The system seeks out its own kind, its mirror images, its chums. It plays (11).

The third great flare is the adolescent sexual maturation. The genital interface explodes on the scene. The system is now chemically and biologically prepared for reproduction.

The fourth lesser flare is then the integration into a "Gestalt" (a view) of comfortable orbit with a sexual partner. It doesn't have to be, but in many species and all mammals, it is the pattern that insures a continuation of the species by protection of the unprepared young system.

These great flaring instabilities all cast light on the orbital configurations that represent a large segment of the behavioral foci. What are the other foci - in particular, the other long-range foci for the complex human animal? The following speculative schema is proposed:

At one or more times within the child-chum-adolescent-young adult stages, there arise particularly favorable orbital configurations - 'experiences' - that are 'attractive' or 'pleasurable' to the individual. A regular pattern grouping of internal and external oscillators occurs - neural, sensory, organistic, motor chain - that are attractive and non-anxiety producing. These make an impression. They are learned as analogs. The system forms preferred paths.

Within a complex of possible patterns, the motions gradually become smoother, more practiced; the system encompasses these into a more determinate pattern of orbital paths that then make up the 'life' postures. The system begins to lock into a more permanent, more characteristic pattern. These paths circulate around the foci of the system; i.e., there exists a focal imperative. The system does not drift through life aimlessly. It is unstable. It seeks to entrain, it selects both foci of behavior and orbits. These 'suit' the individual. They involve more and more routines that become subcortical. They begin to form a patterned field in the brain. Socially patterned foci emerge. Man becomes doctor, drunkard, woman-chaser, intellectual, politician.

A fundamental observation to make about the content of behavior is that externally it does not appear to have a strong metric, whereas internally it may have. What this proposed principle means is that, internally, in the hierarchy of oscillator chains, there is a parity of measure by which the priority of effects and probable sequencing[1] of oscillator states has some

[1] The sequencing is more diffusive than rigorously wave-like, and more probabilistic like a Markov chain than rigidly deterministic.

deterministic but weak connectivity. Externally, the response to input sequences is 'irrational', i.e., not highly ordered in regular sequence, for the individual. He may get mad at a trivial input; he may disregard obviously near-infinite forces; he will run uphill; adore the indifferent; be swept into step by the outrageous; mistrust his most certain guides. His 'tastes' and 'fancies' have traditionally defied accountability. Yet, internally, there is a psycho-logic of sorts. We thereby most often know our man.

Here then is the central theme of a man's behavior - a choice of his life's focal pattern that fits or becomes part of his slowly changing image ideal.

In the present context, the image ideal is a very primitive integrated pattern of internal operative states, involving the many body images that are projected into the brain, which provides a satisfying state for the total organism (e.g., involving surface temperature, pressure, sound field, light field, interpersonal constellations, etc.). 'Satisfying' states are those that were likely associated with early satisfying physiological experiences. The image ideal is not a 'valued' pattern, but an abstract pattern of internal physical-chemical states.

It emerges in essential form in man by age three. It develops for the first two years of life within the mother-child constellation (the toddler stage), and during the third year within the mother-father-child constellation (as the child becomes an active motor system). This image ideal pattern lasts through all life. Its subsequent more socialized development simply represents a veneer on the basic structure. The basic structure remains the patterned response in the endocrine system that the nervous system has been adapted to produce.

Longer term behavior may be represented by the continued effort to shift the internal state patterns so as to attempt to bring them into concordance with the image ideal.

III. HIERARCHICAL REGULATION IN THE COMPLEX BIOLOGICAL ORGANISM

Having furnished an introduction to how the biological 'factory' organizes its activities in some general fashion, it is desirable to furnish an introduction to the hierarchically organized complexes that are found within man.

1. Biophysical Preliminaries

Recapitulating our physiologically founded view of the biological organism, its understanding may be organized around the following levels:

(a) biochemistry at the molecular level;

(b) process maintenance and exchange at the cellular level;

(c) process maintenance and exchange at the organized level of organs;

(d) internal process organization and logic for the macrosystem (i.e., overall systems analysis);

(e) 'factory' operation of the biosystem in its total environment (i.e., both internal and external dynamic systems behavior);

(f) genetic and epigenetic coding for reliable reproduction of biosystems that are operative in total ecological environments.

We select the exposition of levels c-e for our first long term task. The previous sections, predominantly limited to organ and total system levels, have led us to the need for formulating a second position statement for the level of factory operation.

Preferred tools for analysis of a gross system and its major subcomponents, studied at that structural-hierarchical level, are its dynamic responses both steady state and transient. This involves the determination of the time and spatial course of the major identified fluxes and potentials, and the modes of motor actions that the system exhibits. (We call this process of analysis 'biospectroscopy', since it relates to physical-chemical spectroscopy.) If these are well identified, then the 'chains' of causality, as these fluxes and modes course through the system, may perhaps be identified by physical-chemical hypothesis.

The two great logical divisions of the nature of flux (i.e., flow) are the separations into power and information. While modern electrical engineering may favor the information fluxes, a more classically oriented physics might start by examining the power fluxes.

Operationally, we have defined the human, our prototype complex biosystem, as a self-actuated motor (automotive) system that intermittently roams through

its physical environment in search of food. Its principle dynamic properties are that it hungers, feeds, and moves about so that it can continue to hunger, feed, and move about. At the right unfolding time, it couples and reproduces so that the newly formed unit can hunger, feed, and move about. This implies that by his fundamental nature, man involves one more more major internal thermodynamic engines, and engine cycles (for motor capability). Such study involves straightforward power engineering. (Fuel is taken in, energy is liberated and made available for work.)

As a first choice for examination, explorations of the dynamics of the metabolic processes have indicated a ubiquity of internal biochemical oscillators (12).

Finding a near stationary spectrum for the many biochemical parameters that were explored (although the cycles warbled or 'wowed' in frequency with considerable noise, their power spectra indicated large amounts of energy tied up in these spectral regions), we may formally identify the chains in which they are involved as nonlinear limit cycles, everbeating, with internally lossy mechanisms, independent of the starting conditions. Since it seems clear that the mean states of these variables are those regulated parameters identified with homeostasis, the central concept of the constancy of the internal environment independent of external change, we have proposed its modification. For such dynamic regulation of the mean state, we chose the modified name homeokinesis, to denote mediation, mainly by inhibition or release from inhibition, of a manifold of oscillatory processes which make up the many biochemical chains in the organism.

Having found that thermodynamic consistency exists in describing some of the basic metabolic processes, and being able to postulate causal chains involving hormone interaction for a number of those identified, we have proceeded to the behavioral logic of the entire system. The biosystem is not a mindless thermodynamic engine system that tracks a routine path, doing its 'thing' over and over again. The system is essentially marginally unstable, and exhibits indeterminate gain at zero frequency.

Roughly speaking this means that if the system is put down it will not stay at rest indefinitely. If in a disturbed state, it will calm down, in time, when put into a confined region. To completely describe its character, we have to postulate a large number of operational modalities in the biosystem. Because of its inherent instability, the biosystem threads these modalities, its 'hungers' to form a satisfactory pattern. This was discussed in the previous section, a proposed extension of physiological to include behavioral homeokinesis.

However, this scheme has still not been sufficiently elaborated to provide a theoretical foundation that can fully stabilize the unstable biosystem. We, therefore, have to postulate a hierarchical nature to the total system's behavioral regulation. Regulation is achieved by overlay of system upon system. A preliminary exposition of this thesis as applied to temperature regulation is in (13). A second description is contained in the subsequent section

on cardiovascular system regulation. However, from these more specialized essays, we have caught a glimpse of the overall scheme, which we propose to discuss in this section.

2. Dynamic Limit Cycles in Mammals - The Foundation for A Belief in Homeokinesis

To prevent these dynamic schemes from being viewed simply as speculation, a quantity of spectral data have been searched out and will soon appear in assembled form (12). Also see Richter (14) and Jenner (15).[1] A common feature in the oscillator chains uncovered is a relatively slow cycle, suggesting that their dynamics are not single rate governing steps at the cellular level; yet the cycles are quite fast for the amplitudes exhibited (one must visualize that it is large power that these nonlinearly stable oscillators put into transit)[2]. The amplitude ranges tend to be near-normally (Gaussian) distributed but with finite cut-offs. Typically their maximum to minimum power ratio over an observation of many cycles is 5-6 to 1. Thus we are not discussing small changes.

The observations are of ventilation gases, metabolism, thermal power, heart rate, blood gases, blood fuel constituents, blood hormones, capillary red cell flow, water content, weight, sexual activity, circadian rhythms, motor activity, and some of psychic state.

3. Describing Homeokinesis - Dynamic Regulation of the Organism's Interior

We find that when interested scientists finally grasp the fact that there really is a fairly determinate spectrum of repetitive chains in the biosystem[3], with relatively little stochastic signal (a characteristic signal to noise ratio is of the order of 4 to 1), rather than the more commonly expressed idea of 'inherent biological variability', the next question they raise is what is the significance of these 'rhythms'.

[1] The pace for the emergence of good periodic data is quickening. More recent examples are Terman and Terman (16) on a circadian rhythm in behavior; Oswald and Merrington on a 90 minute rhythm in behavior (17); and Hellman et al, on periodic cortisol secretion in man (18).

[2] Illustrating with a few examples, the blood sugar is found to oscillate in less than a minute cycle with $\pm$ 15 mg % amplitude, heat production is found to vary in the body in a 100 second cycle with a two to one change in power level, in some people activity patterns change drastically within a few days or a number of weeks cycles.

[3] The cycles are not individually stationary, but wander over essentially a stationary stochastic distribution. One might say that the spectral lines are not 'monochromatic' (i.e., exhibiting a single fixed frequency), but instead 'warble'.

We view these rhythms as being at the foundation for regulation. Their up and down variation represents the mechanisms by which the regulation of the mean state takes place. As a brief restatement of their significance: There is a matrix whose columns represent the chemical materials, chemical process chains, and behavioral modalities of the system, and whose rows represent ordered time scales. All intersections are not filled. However, in any column, the many temporal levels of regulation of that material, process, or modality are catalogued. In any row, the elements of material, process, or modality may or may not be coupled. The coincidence of time scales, where processes are not coupled, may be due to the use of similar or comparable physical-chemical steps or spatial domains, so that the rate governing processes are similar. We have indicated that the pertinent frequency range is quite broad, approximately 10 decades.

It is essential then to recognize that the system and its many chains are marginally unstable. It is this character that keeps the system and its internal fluxes in intermittent motion, always in process of responding to the changing inputs to the system, to the vicissitudes of the milieu which appear as an impulsive spectrum.

The net resultant of this instability and the chains of organization that emerge both genetically and epigenetically, are that the system locks nonlinearly into behavioral modalities, perhaps 10-20 in all. (Ethologists identify 10, we count 20. See, for example, Scott (19).)

As a corollary of the system instability and its emergent behavioral modalities, regulation in the complex system is hierarchically organized. The system cannot be viewed as at an indefinite quiescence from which its regulation emerges. It truly has indeterminate gain at zero frequency, in which the steady state of 'rest' has many time scales. Instead, it is in overlay of system upon system by which total regulation takes place. Reference (13) attempted to make this clear with regard to thermoregulation. One must always think of system's regulation as taking place against a total scheduled background of normal activities that make up a standardizable behavioral pattern for the biosystem. (Detailed moment by moment, hourly, daily, weekly, monthly, yearly activity, and finally life pattern.)

4. Hierarchical Regulation

Let us view the operation of the complex biosystem via an analogue. As explorers in a strange land, we note complex 'factories' in operation. We don't know their logic, the 'purpose' of these systems. However, we infer that the ensemble is ergodic. What any one individual 'factory' does resembles what others do. Each individual seems to repeat epochs of much the same behavior. The 'factories' have a start up, a life, and a degradative phase, and appear to have internal loss mechanisms. From this we infer that the observed 'factories' are systems.[1] This requires qualification by the first

[1] Systems to us, have long autonomous lives, start ups in some crucible or factory, and ultimately degradation to failure.

and second laws of thermodynamics, i.e., in the system, energy is neither created nor destroyed; the system goes through an energetic cycle, which thus is the minimum period for thermodynamic equilibrium; as a result of the autonomous processes an irreversible production of entropy takes place.

Next, we determine the frequency response characteristics, of both autonomous oscillatory and transient phenomena. This helps suggest some of the detailed internal mechanisms.

Now we are ready to attack the logic of the system. If we succeed, then we may go on the the 'purpose' of the factory. In the biosystem, we are up to these questions. What is the logic behind its spectra and subsystems?

What we propose to do is outline the temporal-hierarchical levels that are associated with the metabolic and automotive process (namely, the system does eat and move about). Some of the levels are speculations based on a number of observations, others are based on much tighter arguments.

0.1-0.3 second level - detailed motor activities can take place at this scale which are programmed by the brain (e.g., one finds here the highest frequency segments in motor actions). Energetic demands can be put on the system, and individual motor units can put forth energy packets within this time domain (at what is considered to be the reaction time level). A first behavioral response shows itself at the highest frequency for which motor activity can be changed in local regions.

5-20 seconds level - storage capacitance of oxygen as a blood gas provides this magnitude of time constant. At this level, a first small behavioral modality begins to arise. The body tends to shift its postures. (Noted in a number of mammalian species. Contrarily, birds, illustratively gulls, change posture more frequently. It is suggestive that the behavior is a visual-neocortical response.)

30-120 seconds level - fuel oscillations in the blood, and a cyclic supply of red cells through small capillaries (i.e., 2-5 microns in mammals) take place with cycles of this periodic magnitude (20). The red cell flow in these capillaries seems to be independent of plasma flow at this time scale, although in arterioles, which supply many capillaries, red cells and plasma flow are proportional (uniform hematocrit). Most recently, the production of hemoglobin has been shown to be synthesized in this time scale. Thus, overall equilibrium in the system within 100 seconds cannot be assumed, i.e., a full cycle of chemical supply for motor performance has not taken place in less time. Yet the organism can perform a complex behavioral task and leave an oxygen debt whose repayment follows with a time constant of this order of magnitude.

400 seconds level - this seems to be the first overall systems level. (An illustration is offered from Brouha (21) and Costell (22). In starting up a long term task involving the entire set of the body, the time constant for increased oxygen consumption is the previous 100 seconds. On the other hand, the time constant for increased heart rate support is about 7 minutes.

Similar observations in a number of different metabolic systems led us to appreciate the general nature of Brouha's results that blood flow in the cardiovascular system is not totally mobilized in less than 7 minutes.) It apparently arises from the integrative response of the hypothalamus in redistributing the systemic blood circulations to satisfy a hypothalamic algorithm.

Major overall shifts to 'stress', i.e., to overall cardiovascular support for a changed activity, take place with this time constant (e.g., skin vascularity, hypothalamic thermoregulation response). There likely are one or more hormones, involved with local organ systems, that help in the blood redistribution with this time constant. However, this autonomic level is not sufficient by itself to run the system. With only this level available, the system would behave as if it were in a coma (i.e., it is not sufficient that cardiovascular system and local organ hormones provide supportive follower action. The system's total actions are not thereby regulated).

20-90 minutes level[1] - because of the large buffer capacitance provided by carbonate, this is likely the epoch of CO_2 equilibrium in the body. Within this time range, it is possible that some CNS segment tends to drive the body toward motor deactivation or activation, in which the system shifts its action via major behavioral modalities. Specifically, at this time scale, we suggest that CO_2 makes the system stir toward a different activity. (Twenty minutes of sustained activity probes at the overall body commitment. It is difficult to concentrate on one thing for longer periods when in a warm, humid, or stale atmosphere, as compared to a fresh breezy atmosphere low in CO_2.)

3½ hours - the overall body thermal balance takes place at this time scale, i.e., it is the primary thermodynamic time constant of the system. The system requires an extensive integrative communicational 'vote' from the various subsystems whether to persist in accepting the behavioral pattern of modalities or to change to something new. (At this time constant, the body makes a volitional, i.e., computed, decision whether to persist in a state or not; whether to work for the next four hours, whether to take an extensive walk in the cold, etc. It requires a complete body commitment. It is speculative, but it is likely the reticular core - as modelled by Kilmer and McCulloch - that acts as a longer term probabilistic computer for one or more epochs lying in the 80-200 minute range. As a further speculation, it is quite possible that some endocrine chain is involved.[2])

[1] In this range, as in many of the other ranges, we are far from certain as to how many spectral lines of independent causality may lie within the stated bands. Like in the early days of electro-magnetic spectroscopy, considerable effort will be expended in uncovering the 'fine line spectrum' of biospectroscopy. Thus we have extended the range to include the 90 minute epoch of REM which has recently been extended to waking behavior ((17), also see p. 79 (23)); but we certainly cannot assert its causality.

[2] Recent data (18) suggests the possibility of cortisol.

24 hours - the circadian epoch involves the automatic making of large scale decisions - with regard to rest-wake, water balance, ingestion. A major patterning cycle of behavior takes place. The complete cycle of performance of the higher nervous command-control system emerges for the first time. Behavior is patterned into rest, wake, eliminate, forage, etc., back to sleep and rest. Corticosteroids and a number of other hormones, as well as the reticular core and other major CNS compartments are involved.

3½-7 days - both physiological and social behavioral rhythms emerge, mediated by hormones with longer action times. A water cycle exists, namely, weight rises and falls with a rhythm in this range. Such cyclic changes in total body water occur via cyclic imbalances in intake and excretion. This quite possibly influences the food and social cycle. (For example, it is approximately the rebalance period required for reacclimatization after time or climatic zone shifts. One may suspect that these readjustments involve a water escapement.) The social week, with its complex behavioral rhythm, fits physiological cues. (Would men, free roaming in the field or within a loosely coupled democratic society, accept very rigid daily schedules without weekly or longer variation?)

20-60 days behavioral rhythms - we believe that a case exists, centered on the menstrual cycle cue, for major behavioral rhythms at this scale. Anxiety-euphoria, intellectual, sexual, 'creative', 'learning' phenomena, it is proposed, possess such periodic epochs. (Taking on a total 'postural' frame of CNS reference - such as intense grief, high excitement, intense activity preoccupation, a complete change of pace of a vacation - can only be held for time scales of this magnitude.)

1 year rhythms - the seasonal cues drive the complex animal, from rigid seasonal movements in birds and fish to smoother ones in mammals. One can suspect that the mammal, such as the higher primate, who exists in a one climate environment - whether hot or cold - does not show much color in his behavioral patterning. His creativity is more limited; his depths and heights - of aggression; of sex activity; of ranging patterns, etc., seem to be less.

5-20 years - the life epoch, the behavioral foci of life. Men exhibit such time scales in their behavioral patterns.

Metabolic regulation takes place with regard to all of these levels. However, the first integrative level for metabolism is of the order of 400 seconds.

5. Applying the Concept of Hierarchical Regulation

It may help clarify the concept of a hierarchy among the overall patterns of behavioral regulation by further illustration.

Can we prescribe an 'average' activity for a person which will resemble the average behavior of a human being? It will neither be rest nor high

activity. There are many body postures possible - rest, eating, motor activity, etc. We see these as physically cast atomistic fragments of the behavioral modalities. However, the time averaged metabolism over some such moderate behavioral activity patterns sustained over a few such time constants (i.e., approximately 20 minutes or longer) is not very dissimilar from the time averaged metabolism over an entire day, an entire weather season, or a 20 year epoch. Thus we can conceive of a systems description of time averaged metabolism over such 400 or more seconds epochs. This measure, taken for some standardized motor activity, could be regarded as an elementary 'normal' metabolism, rather than a 'basal' metabolism.

However, let us not look at the problem as the determination of such an average or characteristic metabolic event. We want instead to know what is a dynamic physiological 'steady state' in which the regulatory functions of interest (here metabolic, thermoregulatory, activity, and behavioral) are implied, i.e., what is it that describes the system moving to eat so that it can continue to move to eat. Such a state will be patterned rhythmically around the average metabolism.

Such regulation does not take place at one level, nor is it rigid. As homeokinesis, it is a comfortable threading of all the human (or mammalian) hungers by small integral numbers per relevant epoch, so that all systems are kept near their regulated state. This requires chains with both internal and external links, even for the apparently 'pure' physiological ones.

Thus there are many normal physiological patterns. All we can do for test purposes, if we want to characterize the system's ergodic character, is to choose an artificial one that samples all of the hungers.

Let us arrive at it by a personalized dialogue: Will you be someone's body slave? More palatably, will you (as employee) be someone else's (as employer) body slave? The answer would seem yes, for a few minutes or a 5 hour experiment. It would appear to be more dubious for a rigid 8 hour, 24 hour, or 30 day experiment.

For each increasing scale of time, the employer must permit more accommodative schedules. For example, up to five hours, he can request a nearly pure single activity. (Lie quietly on a bed, sit quietly in a comfortable seat, exercise steadily at moderate rate, eat slowly but steadily, enter into a sustained argument, look at pictures.) For 8 hours, the subject would want a food break and an elimination break. For 24 hours, the subject might request rest-wake, a few food breaks, one or more elimination breaks, a number of no attention (i.e., relaxation) breaks, and a few breaks for interpersonal attentions. For two weeks, the subject would want these and a few more, to be available on a repetitive, near circadian rate. However, he may be willing (barely) to accept them as a very rigid routine. For 30 days, he is no longer willing to accept these schedules as a rigid daily routine. Yet at the time scale of a month, by introducing a few periods (at least hours, perhaps days) of more patterned threading of modalities, a 'normal' life is feasible.

We have proposed the list in the previous section as necessary modalities in the human. We will not assert their independence (they are certainly not 'normal modes' in the physical sense), nor their certainty, nor their uniqueness. Yet, intuitively, even after comparing with others which we have subsequently found, we still believe that the list furnishes nearly an upper bound to the number of salient mammalian modalities. A more unique identification of modalities may take a full generation to study.

Thus an employer does in fact tie an employee up for years; the entrepreneur is in fact also tied by his boss, the customer, up for years. Loosely speaking, in that time scale, a man is willing to pattern his life almost on a rigid yearly schedule with a minimum amount of variation (e.g., comparable to the 4 to 1 signal to noise ratio referenced before.)

Whether incarceration can deal with man or animals on more rigid schedules needs considerable looking into. Laboratory animals are quite different from wild animals; they require a large measure of support of function for their maintenance; and their schedule is less rigid than might appear.

With one further illustration, we can attempt to bring the hierarchical nature of behavioral regulation even more clearly into focus. The subject in this case will be how the thermoregulation of an individual is achieved through his behavior and in his branching choices in life.

At the 10 Hz. level, his individual muscle fibers and nerve fibers can be used for very fragmentary actions. (For example, he may play a musical instrument. If musically competent, he can step the tempo up to play nearly 10 notes per second, involving individual motor actions (24).)

However, his blood follower system must power this activity. Thus a complex of metabolic blood followers will be found cycling away at the 30-120 seconds level to provide fuel, oxygen, and the hormone drive to sustain that activity. As evidence for this, if he stops the activity, a metabolic decline can be shown with about a 100 seconds time constant. Alternately he can (and often does) conduct his system to nearly any all-out activity for a few minutes. (Almost anyone can be trapped to run around a gym; swim two laps, etc. under ordinary circumstances.) Heat pulsing from the body can be found at this time scale.

However, if one wishes to sustain his new activity pattern, he can no longer do this on anerobic storage. The power conversion must be aerobic. (For men, this likely represents an oxygen consumption of less than 3 l.p.m.) The body follower system must drag the supporting cardiovascular system along. It is here where the hypothalamus comes in, particularly its thermoregulatory response.

It is our best hypothesis that the hypothalamic algorithm involves putting the 'fires' of temperature out. We believe that there is a blood division among the peripheral systemic circulations to increase the flow to reduce the local temperature. Thus demands put on systems - the skin heat exchanger, the

GI tract, the kidney, etc. - result in an increased flow to that zone. It is a zonal control of blood flow. We find evidence for the seven minute temperature cycle from Benzinger's hypothalamic measurements (25); from Brouha's heart rate rise with activity (21); from Costell's report of rise in heart rate for long distance running (22); from the time constant for pooling of fluids at the extremities; from temperature changes seen in whole body thermography. We find, upon examining Fusco's data (26), that it is this thermoregulatory response which is abolished for weeks with hypothalamic lesions.

It is this level that we believe is controlled by nervous action at the level of arteriolar sphincters, or the muscles regulating small vessel flow. As sustained body demand changes, there is a cycling through the body zones, with whatever system having the major duty getting an augmented blood supply, controlled by arterioles actuated from the hypothalamic level.

This is what keeps the individual thermoregulated at the 7 minutes level. Namely, the hypothalamus has arranged a blood subdivision commensurate with power and oxygen demand and its central regulated temperature.

Will the individual persist in this task? No, he may get bored in 20 minutes.[1] Here there is a 'vote' from the body as a whole, likely weighted heavily by the CO_2 and pH status of the body, as the resultant appears at a coordinating 'center', which reports the result to the reticular formation. Our scheme is, of course, somewhat hypothetical, but it has been adapted to fit a near 20 minute CO_2 time constant.

If the individual works in a CO_2 laden, stale, warm, humid atmosphere, he gets sleepy and bored. If a speaker is not colorful and doesn't know how to stir the listener's internal faster hormones, the listener's attention departs (typically, the speaker can use adventure, danger, or sex in his talk, but not tons of data, abstract equations, or too many ideas). Conversely, fresh air, a change in pace or activity will change his status.

Thus integrating the response of the various systems that report to the reticular core, a 'vote' is taken to continue the particular state of attention or change the activity level. One of the signals that is pressing is the temperature regulating response. After the 7 minutes response, the individual's extremities go down in temperature, if it is cold. The systems cast their vote whether to persist in the activity or not. (It is such an explanation, depicted as a probabilistic command-control computer, which Kilmer and McCulloch have been attempting to model in the reticular core.) It is only then the extremities that might be signaling discomfort. Yet the person can make the overriding decision for his body to continue or to change the activity.

[1] As we stated before, we cannot distinguish, except loosely, between a possible 20 minute CO_2 epoch and a 90 minute waking equivalent to REM. Both seem to arise from coupling at the level of higher nervous activity.

The body algorithm is ruthless. If you want to persist, it says, you may, but it will no longer waste the heat (Algorithm: if a region is warm - give it blood, if it's cold - take it away). The peripheral circulation to the extremities begin to shut off. The large body follower begins to make large thermal adjustments. Behind it stands the overall chemical balance in which the CO_2 and pH system and its related hormones likely dominate.

The internal systems of the person are thus capable of voting locally, i.e., major body zones push up their satisfaction-discomfort level at higher centers, and a vote is taken to persist or desist.

Now for a more integrative level. Consider undertaking a time consuming task such as a 3-4 hour walk in the cool damp woods.

Note at this level there is already a near conscious vote involving all of the systems and the memory, making use of remembrances of things past. There is an estimate of the level of physical discomfort, the possible level of joy and satisfaction. (Choosing depends on trade-offs of goods and evils.) The time constant here is a full thermal time constant of the entire body. It is also, incidentally, the approximate relaxation time constant for food satiety.

If the person commits himself to such long term actions the entire body thermal equilibrium will shift. Within limits the system is capable of readjustment of the heat exchange blood flow that will keep its hypothalamus temperature regulated (not controlled - Benzinger's data (25) show there can be a near 1°C shift in level). There is a corollary regulation curve for metabolism.[1] There is a range of environment and sustained activity over which the body can perform. It is a wide range, but not infinite.

Of course, the human has augmentors, extensions, such as clothing. However, nearly nude, he can tolerate 0 to 40°C by activity regulation.

In longer term, the person cannot sustain the activity. Day-night enters. His reticular core finally votes and tells him to turn off the light. He sleeps.

With regard to a circadian temperature cycle (13), it is not clear whether there is an autonomous thermal cycle independent of an activity cycle. We believe there may exist a rest-wake-activity cycle (food foraging, etc.) and that temperature changes are concomitant with the activity changes. However there are studies to contradict this belief (28).

[1] It costs more to perform a given task at extremes of cold and warm. When the body is well soaked at a given temperature (e.g., 3-4 hours), the average metabolism has been shown to be a broad U-shaped, somewhat poorly regulated function of ambient temperature at a given activity level (13). A similar result is found with the body immersed in water (27).

Note that the experimental situation of clamping a person in a fixed modality; or actually in a pattern involving very minimal modalities is almost an impossible task to perform. This is the essential meaning of the statement that the system has indeterminate gain at zero frequency (i.e., at frequencies smaller than 1 cycle per 12 hours or 1 cycle per 24 hours). The problem is even more aggravated with sensory deprivation, as more modalities are excluded. (Namely, there is doubt whether a total sensory deprivation experiment could be conducted for 24 hours.)

Thus it can be suspected that temperature regulation at the 24 hour level may be a concomitant of the rest-wake pattern of life. A test of the truth or falsity of this assertion must be the subject of carefully designed experimental verification.

At the next level, while there is a strong social basis for weekly patterns (e.g.,the Thank Goodness It's Friday syndrome), a strong 3½ day water cycle in humans and in guinea pigs can be found. The data available have not completely divorced the cycle from the social week. Nevertheless, they are suggestive that the time constant for water is of the order of 3½ days, and that the food cycle entwines that water cycle. There is reason to believe that acclimatization, say under movement from one time zone to another, from one temperature zone to another, from one altitude to another, from one heavy activity pattern to another, has a time constant of this magnitude (3-5 days) and that it is associated with the water balance. Thus water and long term metabolism are certainly coupled in one direction, and possibly have nearly symmetric coefficients, coupling them the other way. Thus the longer term thermoregulation is possibly tied to this water-metabolism-activity cycle.

Now we enter the near purely behavioral range of 20-60 days. In this domain, the chains certainly have external links. (In fact, we consider that the distinction between psychological, as opposed to physiological behavior, should be made on the basis of how strongly the chains of causality depend on external links that are directly in the chain, rather than at boundaries.) A major polarization for a male is afforded by a wife's (or involved female's) autonomous menstrual rhythm. Beach (29) points out both female humans' and rats' activity levels vary over the menstrual cycle. (Also see Hafez (19).)

However, it is likely that the female influences the brood within her sphere - husband, boyfriend, and children. Thus a subtle emotional cuing exists. This tends to entrain other emotional derived phenomena.

On the other hand, we believe, with mostly behavioral evidence so far, that there are autonomous emotional, intellectual, sexual, anxiety-euphoria, and other rhythms vaguely in this time domain, that involve definite internal, mainly biochemical chains. However, they have strong external links, i.e., it is a time constant of the order necessary to have a good vacation, get an important idea, work up courage to get married, or to get drunk, to get a new job, to become very anxious, or euphoric, etc. All of these 'volitional' patterns lead to very definite physical-physiological involvement. In social life, the female's menstrual period sits there as a gentle signal rocking the system's

frequency response. Temperature regulation in this domain is more subtle. It is the follower pattern on the activity patterns of a busy life.

Since that temporal level may have seemed a little obscure (for temperature, not for living) the next one may be more clear. We come to the seasonal cue.

The human chooses a thermal environment he can make out in. (One may not overlook his clothing and his environmental thermal conditioning.) On earth, man responds - if he is temperate zone adapted (or even if he is not) - to work out a seasonal roaming. He may go south in the winter, and north in summer.

While the reaction may be intentionally exaggerated for the affluent human, it is clear that many species adapt to the warm-cold of seasonal changes in a great variety of ways. Birds migrate, many mammals change their insulation. Many primates change their environment winter and summer. The complex human tends to control both his locomotion state, his local living milieu, and his location. Thus, in civilized society, he simply shows an even more complex reaction pattern.

However, basically, the human works out motor patterns that hold his thermoregulating signal within bounds. He does not conduct himself with heavy activity in very hot climates, nor does he overexpose himself in very cold climates. At near 0°C with little clothes, he has to have a very active kind of daily and seasonal pattern.

Beyond lies a slow adaption. We suggest that the mammals (suppose them to be driven by changing weather cues, such as a cold season to season spell, or a warm one, or a dry one, etc.) begin to explore the ecological environment, drifting slowly in it, in search-seek patterns which have a strong thermoregulatory cue. This is at the seasonal or longer level. Animals change their roaming habits over years, not necessarily over days. This seems to be the record read from the archeological evidence of past ice ages (e.g., in the wandering of mammalian species over Europe).

Thus the thermoregulatory homeotherm does not achieve his regulation by one level out of a hierarchy (e.g., hypothalamic regulation) but by overlay of level upon level.

While so far we have discussed this for only one signal, with the temperature signal in the foreground, the ideas and modeling are true for all other essential fluxes - for light, food, sex, etc.

In such fashion, we have attempted to illuminate hierarchical homeokinetic regulation with its strong behavioral overtones.

IV. SUMMARY OF SOME OF THE ABSTRACT PRINCIPLES FOR A BIOSYSTEMS SCIENCE

As a summary of this introductory section, we may briefly review the principles thus far proposed as fundamental to a biosystems science.

1. Homeokinesis - a description of biosystems dynamics. - A realization of Cannon's principle of homeostasis encompassed in the recognition that all internal systems in the biological system consist of limit cycle oscillators, and that the system is governed both chemically and electrically be mediating the stability of these oscillators.

2. Spatial and temporal biospectroscopy. - There is a matrix whose columns represent the chemical materials of the living system, the chemical process chains of the living system, and the behavioral modalities of the system; and whose rows represent ordered time scales (As illustration, see the figure). Not all intersections are filled. However, in any column (e.g., one labeled water, or pH, or ingestion), the many temporal levels of regulation of that material, process, or modality are categorized. In any row, the elements may or may not be coupled. The coincidence of time scales, where processes are not coupled, may be due to the use of similar or comparable physical-chemical steps or spatial domains, so that the rate governing processes are similar.

3. System instability. - The characteristic property of the system and its many causal chains is its marginal instability. It is this character that keeps the system and its fluxes in intermittent motion, always capable or responding to the changing vicissitudes of the milieu.

4. Behavioral modalities. - The net resultant of this instability and of the chains of organization, that emerges both genetically and epigenetically, is that the system locks into behavioral modalities, perhaps 10-20 in all.

5. Hierarchical regulation. - As a corollary of the instability of the system and its emergent behavioral modalities, regulation in the complex system is many-leveled. It is not the case that the system can be viewed as at an indefinite quiescence from which the regulatory system or systems emerge. The system has indeterminate gain at zero frequency. Instead, it is an overlay of system upon system by which the total regulation takes place. The previous section attempted to make this clear with regard to thermoregulation. One must always think of systems' regulation as taking place against a total scheduled background of the normal activities that make up a standardizable behavioral pattern for the biosystem. (Detailed moment by moment activity, hourly, daily, weekly, monthly, yearly, and finally life pattern.)

In order to probe at these major theses, the argument that had been heretofore conducted was to sustain a broad discussion of all of the system chains, ever seeking pertinent experimental data, but to pursue the exposition of a few specific chains (see, for example, (30) for a more detailed statement about many of these chains). In particular, considerable effort was spent in the exploration of the near 100 second time scale. The periodic (limit cycle)

Time scale-- sec.	Materials: Atomic species: O, C, H, N, Ca, P, Na, Cl, Fe, S, K. Molecular species:																
	H_2O	CO_2	O_2	Protein			Structural Calc.	Fats	Ions	Carbohydrates			Heme	Urea	Nucleic acids	Minor Constituents	
.03-.1				Structural Protein	Enzymes	Hormones	(Carbonate, phosphate)			Glucose	Glycogen	Glucophosphate				Hormones (Steroid, Phenol)	Vitamins
.1-.3																	
.3-1																	
1-3																	
3-10																	
10-30																	
30-100																	
100-300																	
300-1000																	
1000-3000																	

Processes

Exs: $O_2 + HbCO_2 \rightleftharpoons CO_2 + HbO_2$ | $C_6H_{12}O_6 + O_2 \rightarrow CO_2 + H_2O'$ | n-amino acids → protein.

Modalities

Exs: Ingestion | Food Search | Sleep | .....

Illustrating the Temporal Matrix for a Given Living Species (not complete)

14. Richter, C., "Biological Clocks in Medicine and Psychiatry," Thomas, 1965; in R. Michael (ed.) "Endocrinology and Human Behavior," Oxford, 1968.

15. Jenner, F., in "International Reviews of Neurobiology," Vol. II, Academic 1968.

16. Terman, M., Terman, J., Circadian Rhythm of Brain Self-stimulation Behavior, Science, 168, 1242, 1970.

17. Oswald, I., Merrington, J., Cyclical "On Demand" Oral Intake by Adults, Nature, 225, 959, 1970.

18. Hellman, C.L., et al., Cortisol is Secreted Episodically by Normal Man, J. Clin. Endoc. Metab., 30, (4), 411, 1970.

19. Scott, J., in E. Hafez (ed.) "The Behavior of Domestic Animals," Williams and Williams, 1962.
in E. Toback (ed.) Experimental Approaches to the Study of Emotional Behavior, Ann. N.Y. Acad. Sci., 159, 1969.

20. Cardon, S.Z., Oestermeyer, C.F., Bloch, E.H., Effect of Oxygen on Cyclic Red Blood Cell Flow in Unanesthetized Mammalian Striated Muscle as Determined by Microscopy, Microvascular Research, 2, 67, 1970.

21. Brouha, L., ASME Paper 57-A-55, 1957; Adv. Mngmt., 19, 9, 1954.

22. Costell, D., Metabolic Responses during Distance Running, J. Applied Physiol., 28, (3), 251, 1970.

23. Iberall, A., Ehrenberg, M., Cardon, S., General Dynamics of the Physical-Chemical Systems in Mammals, Aug. 1966 Report to NASA; Gen. Tech. Serv., Inc., Upper Darby, Pa.

24. Smith, H., "From Fish to Philosopher," Ciba Ed., 1959.

Quastler, H., Three Survey Papers, U. of Ill. Control Systems Lab., Report No. R-71, 1956.

25. Benzinger, T., Kitzinger in "Temperature - Its Measurement and Control in Science and Industry," Vol. 3, Pt. 3, Reinhold, 1963.

26. Fusco, M., Ph.D. Thesis, U. of Penn., 1959.

27. Craig, A., Dvorak, M., Thermal Regulation of Man Exercising During Water Immersion, J. App. Physiol., 25, (1), 28, 1968.

28. Colin, J., Timbal, J., Boutelier, C., Houdas, Y., Siffre, M., Rhythm of the Rectal Temperature during a 6-month Free-running Experiment, J. App. Physiol., 25, (2), 170, 1968.

29. Beach, F., "Hormones in Behavior," Hoeber, 1949.

30. Iberall, A., NASA Contractors Report CR-129, Oct. 1964.

a value near the optimal pH for its uncoupling action. Mitchell (13) postulated that oxidative phosphorylation in the respiratory chain causes a pH gradient to be developed. This transfer of H^+ or OH^- is the driving force for the phosphorylation of ADP. Furthermore, he proposes that the uncoupling agents act as pH "sinks". That is, they do not interfere with the enzymatic activity along the cytochrome chain, but prevent the formation of the pH gradient by acting as lipid soluble proton carriers. This theory corresponds to the observed characteristics found in the bilayers.

In addition, the permeability sequence of bilayers to cations ($H^+ > Rb^+ > K^+ > Cs^+ > Na^+ > Li^+$) is nearly identical with that for biologic membranes ($H^+ > K^+ > Rb^+ > Cs^+ > Na^+ > Li^+$) (14).

Since the evaluated characteristics of biologic membranes and lipid bilayers are so similar, it is postulated that the structural relationships are also similar. The orientation of the molecules within the bilayers corresponds to the lipid layer in the Davson-Danielli model of the biologic membrane.

A common component of all living membranes is a structural protein. Although the exact nature of this protein has not been ascertained, its monomeric form has been found to have a molecular weight of from 20,000 - 50,000. The type and configuration of these molecules are influenced by genetic coding. Maddy and Malcolm (15, 16) have extracted the protein and found that most of the molecules are in the α or globular configuration, with a small percentage in the β or extended configuration. One theory, proposed by Lehninger (17) is that the nature of the protein coat array determines the nature of the lipid layer. There may exist a specificity between the code for protein and its structural correlate, the membrane lipid.

In addition, the membrane contains carbohydrate adjuncts. Generally these are in the form of mucopolysaccharides, glycolipids and glycoproteins. These saccharides are linked with the protein coat of the membrane. Lehninger (17) points out that the side chains of the oligosaccharides contain sialic acid. These moieties are covalently linked to the asparginine residues in the peptide chain. Other groups may be linked to the serine, threonine and hydroxylysine amino acids. The bulk of the polysaccharide molecule extends out into the intercellular space.

Each component of the membrane contributes to its characteristic physical attributes. The exterior of the membrane has a negative surface charge due to the sialic acid groups attached to its surface. (18, 19, 20).

2. Informational and Functional Substances in Membranes.

Also contained within the surface structure are antigens, enzymes and potential "carriers". The most common of the antigenic molecules are those producing the characteristic blood groups. These are the ABO, MN, Rh and other classified blood typing antigens. These agents react with foreign

antibodies, carried in the extracellular medium (plasma). When in contact, antigen-antibody links are formed. These result in the formation of large conglomerates, or clumps. The antigens are contained within the outer layers of the membrane. Of the enzymes incorporated into the membrane structure, ATP-ase has been studied most extensively. Adenosine triphosphate provides the energy for the transport of sodium ions across the membrane. Other phosphorylated purines or pyrimidines cannot be substituted for adenine. Many characteristics of this sodium pump have been characterized by Hodgkin and Keynes (21).

The enzyme ATP- ase is a lipoprotein, which requires the presence of magnesium ions for activation. Calcium ions inhibit its activity. ATP- ase is oriented across the membrane, such that it is more selective for sodium ions within the membrane and potassium ions on the exterior (22, 23).

Depletion of erythrocyte ATP causes the loss of potassium and the gain of sodium. These effects are accompanied by an initial swelling and later shrinking of the cell. In addition, the viscosity of such cells is increased and their membrane elasticity is decreased. The minimum critical diameter through which a "normal" human erythrocyte can pass in vivo is 2.8 μ.[1] ATP depleted cells can pass through a pore no smaller than 3 μ. (24).

Similarly, depletion of membrane ATP increases the stiffness, such that after 24 hours, 60 mm H_2O pressure is needed to deform the cell. As little as 4 mm H_2O will deform a normal cell. Restoration of the membrane ATP restores the elasticity.

Elevation of the level of calcium ions, an ATP-ase inhibitor, also causes stiffening of the membrane. Removal of the calcium, with a chelating agent, such as EDTA restores normal elasticity.

In addition, Bensch and Bensch (25) have shown that hypoxia causes an increase in binding of ATP and diphospho-glyceric acid to the hemoglobin. This transposition from membrane to hemoglobin removes the ATP from its active membrane configuration. Although facilitating the release of oxygen from hemoglobin, such binding may inhibit the deformation of the membrane.

Other enzymes such as permeases (26), catalases (27, 28) and estereases are also bound within the membrane. The phosphotransferase system (26, 29) which includes the permeases, consists of two enzymes, a low molecular weight protein and phosphoenolpyruvate. Each of these components is bound within the membrane in a discrete functional configuration. It is likely that other enzymes are present within the membrane, acting as intermediates in the transport, activation or deactivation of metabolites. The contribution of these

[1] In vivo studies show erythrocytes passing through 2 μ capillaries.

inclusions to the membrane structure is of indeterminate significance. These molecular sites may modify the membrane symmetry.

3. Theories of Membrane Structure

All current theories of membrane structure are based upon the assumption of the existence of repetitive structural units. This assumption has recently been challenged, as described below.

The classic, previously most widely accepted theory of membrane structure was proposed by Danielli and Davson in 1935 (3). They postulate that the membrane is composed of a bimolecular leaflet of lipid molecules. These are arranged with radial symmetry, such that the hydrophobic molecular ends face toward the center of the membrane, and the hydrophylic ends face out. This bimolecular layer is coated on both surfaces by denatured protein. This protein coat reduces the surface tension of the membrane to a value below one dyne per square centimeter. Robertson (5) confirmed this structural assumption in tissues fixed either with potassium permanganate or osmium tetroxide. The most widely analysed membrane is the myelinated nerve fiber. The nerve-Schwann cell interaction culminates with the formation of the myelin sheath. The observed repetitive ultrastructural nature of myelin, the infolding of the Schwann cell membrane, is the basis for establishing the Davson-Danielli model as a structural reality. The dimensions ascertained by electron microscopy are two dense lines, each about 25 Å thick, seperated by a light area, the lipid bilayer. Myelin, the subject of Robertson's (30, 31), and Schmitt's (32) study, is not a typical membrane. It exhibits sparse protein layering. In addition, the staining technique used, although known to react with proteins, also reacts with the hydrophilic olefinic bonds in the fatty acid residues of the lipid layers. The potassium permanganate and osmium tetroxide stained preparations are not identical. The staining and dehydration of the membranes distorts the structure.

4. Other Concepts of Membrane Structure

The tendency has been to assume that the bimolecular leaflet structure extends around the cell as a coat, of which all components are continuous. The membrane is punctured by pores, through which molecules can pass without penetrating the membrane proper. Sjostrand (33- 36) has argued that the structure actually consists of globular subunits. The initially observed laminar structure, he postulates, is due to the uneven distribution of hydrophilic and hydrophobic protein side chains. The hydrophobic side chains lie facing the adjacent conglomerates, whereas the hydrophilic side chains face the surfaces. Since the hydrophilic moieties are preferentially stained, an apparent laminar structure will be observed.

These postulated globular units are protein. Since these globules are structurally independent of each other, each may be composed of an enzymatic element of a metabolic pathway (e.e., mitochondrial respiratory chain). The spatial conformation, held by the hydrophobic bonding, potentially orients each globule as a link.

Studies of membranes by means of optical rotating dispersion, or circular dichroism have suggested other possible membrane models. Membranes may be seen as mosaics of hydrophobic and hydrophilic patches. The hydrophilic patches may possibly serve as aqueous channels through the membrane. Two new models of membrane structure have been presented by Wallach and Gordon (37) and by Lenard and Singer (38). Both of these models emphasize that the so-called "unit" membrane has multiple subunits, and there may be no universal arrangement of these. In fact, as Korn has stated (39): "...the real possibility exists that membranes may differ from one another to such an extent that they cannot be usefully described by one unifying model."

5. Electron Transport Membrane Systems

Within membranes of mitochondria, electron transfer systems can operate presumably because each globular sub-unit contains sufficient rotational activity to present surface conformation changes between adjacent globules. Green and Perdue (40) and Fernandez-Moran et al.(41, 42) negatively stained beef heart mitochondria; they found that the globular membrane structure seen by Sjöstrand is modified in the inner membrane of mitochondria. The elementary particles of cristae membranes, they propose, are composed of tripartite units.-a base (40x112Å) which is attached to adjacent units, a narrow neck (30x50Å) and a detachable head (80-100Å).

These inner membrane base pieces are composed of 45% structural protein and 55% electron transfer chain protein. The inner membrane contains the cytochrome enzymes, whereas the outer membrane contains the citric acid cycle, fatty acid oxidation, fatty acid elongation, phospholipid synthesis and substrate phosphorylation enzymes. The head piece is a water soluble protein of which ATP-ase activity is the dominant function.

Many metabolic activities have been ascribed to different cellular membrane components. Green and Perdue (40) have summarized these functions in the following list:

Metabolic Functions of Selected Membrane Systems

Plasma Membrane	- glycolysis[1]
Endoplasmic Reticulum	- synthesis of protein, fatty acids, phospholipids
Microvilli	- terminal hydrolysis;proteins, polysaccharides
Mitochondria	- citric cycle, fatty acid oxidation, fatty acid elongation
Chloroplasts	- synthesis of sugar from CO_2 and H_2O

1. Glycolysis has been observed on erythrocyte and yeast cell membranes.

Although glycolysis occurs on the plasma membrane, other substrates must penetrate into the internal cellular matrix. These chemical elements can pass through the membrane walls or pores. The path or paths selected is a function of the physical interactions between the membrane surface and molecular substrate. The configuration, size, electrical charge, solubility and concentration of particles are important determinants of selective transmembrane passage.

The unique nature of absorption, adsorption, and excretion of different cell types accounts for the specificity of cells and the formation of discrete tissues and organs.

6. Ion Pumps in Membranes

Many different pumps exist to move ions across the membrane, against a concentration gradient. These are active processes, requiring energy, in the form of ATP.

The most intensely studied ion pump is the Na-pump. This is energized by ATP and exchanges sodium for potassium ions. These ions are exchanged in a 3:2 ratio. The current proposal is that ATP-ase binds 3 Na ions inside the cell; subsequent phosphorylation by ATP causes a conformational change so that Na^+ ion is transferred to the outside of the membrane and detached; 2 K^+ ions are then bound, carried into the cell and detached. The binding also is accompanied by dephosphorylation of the transporting ATP-ase, which is then ready to bind more Na^+. The excess sodium ions leaks into the cell such that the net exchange is 1:1.

7. Resting Membrane Potentials

Concentration gradients for sodium and potassium ions are developed across the membrane by ion pumps. The concentration gradient forces set up electrogenic forces, caused by the disparity in concentration across the membrane of charged particles, both free and restricted, describable as diffusion potentials. (Note that protein molecules at pH 7.3 have a net negative charge.) Since some charged particles are constrained by pumps, or by permeability barriers, a steady-state can be achieved. This balance of forces results in a transmembrane potential.

Within the membrane there exists a lattice of fixed positive charges. This repels cations and attracts anions in the medium. It is found that anions penetrate into red blood cells 10^6 times faster than do cations. The rapid penetration of anions occurs in spite of the negative surface charge. The flux of cations is also affected by the concentration of hydrogen ions within the cell.

As an example of ionic interaction, the flux of sulfate ions across the membrane was demonstrated by Passow (43) to be modulated by both hydrogen ions and chloride ions in the external medium. As the

positive charged hydrogen ion concentration increased, so did the sulfate flux. As the negatively charged chloride concentration increased, the sulfate flux decreased.

8. Action Potentials

We will not discuss here the theory of the ionic basis of conducted action potentials. The neurons of the nervous system are elaborated so that a disturbance of the potential across the membrane can be conducted, and interpreted by a distal receptor, as an informational signal.

The birefringence pattern of the squid giant axon, a myelinated nerve, or the rabbit abdominal vagus, a nonmyelinated nerve, is characteristic of the molecular conformation of the membrane. When an action potential passes down the membrane (the action potential is generated by a sudden, triggered influx of sodium ions), the axon birefringence increases. With the interior of the squid axon replaced by an artificial physiologic medium, the initial birefringence pattern and the shift in birefringence, as a potential wave passes by, are retained.

This electromagnetic pattern of birefringence follows the potential changes and not the current flow. If choline is substituted for sodium in the medium, no birefringent shift is seen with stimulation. The electromagnetic shift is likely due to the reorientation of charged particles within the membrane during the depolarized phase.

9. Coupling Between Solute and Water Transport

The movement of sodium ions is also bound to the transport of water. Fresh water animals must maintain a high salt concentration in a medium of low tonicity. The sodium pump, acting in two stages across both the external medium-cell and cell-internal medium membranes, moves sodium from outside to the extracellular space (44). In this process sodium is exchanged for hydrogen ions, and chloride for bicarbonate ions. The Na-K exchange across the cell-internal medium membrane can be blocked by the cardiac glycosides. The exchange of Na^+ for H^+ in the form of NH_4^+ across the external medium-cell membrane cannot be inhibited.

10. Other Ion Pumps

Four other types of ion pumps also function in biologic systems. Potassium is actively pumped out of the malphigian tubules of insects and the gut of the silkworm. This movement is independent of the sodium flux. In the gall bladder a sodium water pump is operative, without the generation of any potential gradient (standing gradient hypothesis). Hydrogen ion pumps are active in the gastric mucosa of mammals and in the salivary glands of mollusks (D. gallia secretes 4% H_2SO_4 from its salivary glands).

Chloride ions are actively pumped in acinar cells of the salivary glands and gastric mucosa, in the sweat glands, proximal tubule of the kidney and in

marine teleost gills.

11. Transport Modes and Mechanisms

The active sodium pump is an example of one process for moving substances across the plasma membrane. The term to describe this pump is active transport. In active transport a substance is moved across an interface, against an electrochemical potential gradient, by the expenditure of energy.

The process governing the passage of permeable substances across an interface down a concentration gradient, is diffusion. If the movement of these particles is greater than can be ascribed to simple diffusion the process uses energy for a facilitated diffusion.

Group translocation is the movement of a particle across the membrane such that its original chemical structure is modified during the passage across the interface. Other types of phenomena can be classified as solvent drag, in which the solvent is moved across the interface along with the solute, pinocytosis and phagocytosis, in which the membrane envelopes the substance (liquid in the former, solid in the latter) and vesiculates them within the cytoplasm.

Active transport phenomena can even cause countertransport. The movement of molecules and ions, according to this theory, is dependent upon a rotating molecule as the carrier. In one conformation, a selectively active site on the carrier faces the exterior simultaneously with the exposure of another active site on the interior. When the sites bind with their respective particles, the carrier undergoes a molecular transfiguration. The sites are reversed in position and the attached particles released.

A more elaborate mechanism has been proposed which is basically identical to this system. In it the carrier acts as a rotating disk. Along the periphery of the disk are the active binding sites. The attraction and binding with the exchanged particle in this system are dependent upon a steric fit with the active groups along the carrier periphery. With the sites filled, the disk rotates, and moves the particles through the membrane to the opposite side. Here, they are released and the active site binds with a molecule to be moved in the opposite direction.

In either case the selective formation of a bond between the carrier and particle is affected by the concentration of potentially transported particles on both sides of the membrane.

The selective nature of this exchange process can be illustrated by yeast cells grown on L-sorbose. This sugar is taken up by the cell in the same carrier system that takes up glucose. The uptake of L-sorbose follows typical saturation kinetics.

If a cell saturated with L-sorbose is placed in a medium containing glucose and no L-sorbose, the internal sorbose concentration falls as the

glucose is taken up. The mechanism for the exchange of two particles accounts for this countertransport of L-sorbose.(i.e., of glucose and L-sorbose.)

Since the carrier can accept either glucose or sorbose, the particle in higher concentration (assuming equal carrier affinity) will bind with the active carrier site. Thus, the high internal concentration of the non-metabolized sorbose increases the chance of binding a sorbose molecule to the carrier over the chance of binding a glucose molecule.

Schultz (45) found that sodium ions are coupled to many transintestinal nonionic transport processes. The sodium ions are linked with both amino acids and glucose. Increasing amino acids and glucose at the intestinal mucosal surface increases the exchange rate with sodium ions. This mechanism is blocked by the replacement of Na^+ by Li^+ or K^+. Schultz postulated that the Na^+ is needed for the formation of amino acid complexes. Ouabain acts on this complex in one direction only. Since this agent acts on the serosal surface of the intestine, it blocks amino acid and glucose efflux, not influx. The absorption of amino acids and glucose is coupled with serosal sodium to mucosal sodium ionic concentration ratio. For influx mucosal sodium ions are needed for the coupled exchange reaction. For efflux, serosal sodium ions are needed.

Another postulated carrier mechanism is dependent upon the binding of the external particle with an enzyme found in the membrane. The particle or substrate binds with the enzyme on the outer edge of the membrane. They form an active complex, which passes to the inner surface of the membrane. On the surface the enzyme is converted to another form, releases the bound molecule and binds with a different molecular species. This second bound particle is carried to the outer surface of the membrane, where it is released, and the process recycled.

This enzyme linked system is typical of the permease system previously described. For a more detailed description of postulated transport mechanisms the review of Wilbrandt and Rosenberg (46) should be consulted.

In addition to penetration of the membrane, particles can perhaps exchange through pores. Solvent drag, the movement of an uncharged particle across the membrane, along with the passage of water occurs, through the pores. The nuclear membrane is a complex lamellar structure containing pores of about 500Å diameter. Their fine structure is not clear, although it is evident that communication between the nucleus and cytoplasma can occur through these pores, as well as passage across the extensive interface of the endoplasmic reticulum.

An alternate hypothesis to the theories of membrane carrier activated transport systems has been proposed by Ling (47), Troshin (48) and Harris and Pankerd (49). They postulate that the active phenomena are controlled from binding processes within the cytoplasm. Miller (49) proposes this control to be the action of a gel-like layer lying just below the membrane. The selective binding of particles within the cytoplasm, effectively prevents them from exerting any osmotic force. Thus, the development of a concentration

gradient, such as for sodium, is produced by binding the particles as they penetrate the membrane.

The gel proposal of Miller postulates a semi-rigid cytoplasmic structure. If the cellular medium is structurally organized, such as a lattice of hydrated water molecules, as the initial observations of Bratton (50) indicate, then many of the characteristics ascribed to membrane function may in reality be reflections of an intra-cellular activity.

12. Some Overall Remarks

A description of transport, from a systems' point of view, at the membrane level is still quite fragmentary. Among systems' elements that are potentially of interest, there is transport through the independent cell, e.g., the red blood cell, egg cells, etc., through the nerve cell, and through the capillary wall. At present the nerve cell is not our concern. A late reference of some interest is Nachmansohn's (52). Transport through the cells, and in particular the red cell, is reviewed or discussed in Davson (53); Dick (54); Loewenstein (55); a "Discussion Only" proceedings of the 1966 Biophysics Society meeting (56); Stein (57); Jörnefelt (58); Dowben (59). These are just a few sources that illustrate discussion from recent years of a rapidly growing field of interest.

The major thrust of these sources is that these cellular membranes permit easy passage of water-soluble small molecules, and much more difficult passage of large molecules. On the other hand, lipid soluble molecules pass through easily. (See, for example, Fig. 3.5, Stein (57)).

The lipid soluble molecules, it is believed, pass by solubility diffusion. The small molecules, it is believed, pass as a bulk flow through 'pores' of the order of 4Å in radius. The evidence for fixed pores is moot, so that a variety of models are under discussion. Besides a simple viscous flow, there are various kinds of electrical binding proposed, not only along the 'pore' path, but at surfaces or within the volume of fluid within the cell.

Transport through the capillary wall is discussed in Davson (53); Zweifach (60), and Bruns and Palade (61).

Large molecules, it is believed, pass via vesicles showing a form with the order of 300Å neck and 600-800Å near spherical body. Small molecules pass through the equivalent of 40Å radius pores. However, evidence for fixed pores is quite weak.

Beyond the problems of a 'passive' but complex diffusion through membranes, whether by solution, fixed or variable pores, or other specific mechanisms, there is the 'facilitated' diffusion which requires active thermo-

dynamic 'engines'.[1] The problem of whether the processes are active or passive (even 'passive' transport coefficients such as 'resistance' - whether electrical, thermal, hydrodynamic, or osmotic - simply relate to active processes that take place at a lower hierarchical level) lies in the identification of the spatial and temporal spectrum associated with these processes. While the electron microscope has begun to resolve the spatial problems - with admitted problems - the temporal spectrum is still in its infancy.

On another front, intercellular communication, as exemplified in the work of Loewenstein (55, 62) clearly suggests that an active calcium 'pump' permits the binding of cell to cell, and the passage of both electrical signals and molecular transport by various sized molecules via the short circuited common membrane junction.

It is clear that from the lower level of molecular biology, these developments in the response of the cellular level or the capillary segment of endothelial cells form a murky but tantalizing integration of molecular processes. Thus they are suitable topics for the molecular biologist but are beyond our current scope.

13. Summary

Ultimately, we desire a synthesis between supramolecular (molecular aggregate) biology and the transport phenomenology described above. At the present time some attempt has been made to rationalize the phenomenological coefficients for water transport through cell membranes in terms of frictional coefficients at a lower structural level. However, the idealized models do not fit recent data simply (63). Permeability coefficients for solvents, that should be constant under specified conditions according to available theory, turn out to be functions of solvent fluxes. Again it becomes evident that membranes are not simply inert, rigid filters, with right cylindrical pores and paths for water.

In the future we forsee that even processes that are regarded as "passive" by the biologist today, will be revealed as depending upon chemical processes in membranes, occurring rapidly, giving the membrane dynamic lability.

14. Addendum - A Speculative Note on a Dynamical Hierarchy of Active Transport Systems

Biological systems are usually represented by a structural hierarchy consisting of molecules, organelles, cells, tissues, organs, etc. The notion of a dynamical hierarchy as we have pursued it here, consisting of both spatial and temporal features of organization, is not commonly advanced as a basic aspect of living systems. We consider the notion at this point, not for novelty, but because it is essential

[1] It is gratifying to find Stein - in discussing active and passive transport - also utilizing such language. He states that "...an electrochemical gradient is a source of energy - it can be used to drive a chemi-osmotic engine...".

to an understanding of transport processes in living systems.

To begin, we note that the terms hierarchically macroscopic and microscopic can be given more precise dynamical meanings by introducing the time-averaging operation $\langle\ \rangle$. If Q denotes some time-dependent physiological quantity then $\langle Q \rangle$ denotes the dynamically macroscopic quantity under consideration at some hierarchical level corresponding to the microscopic quantity Q at a lower level. Dynamically macroscopic quantities occur naturally in the description of most physiological functions because of the various relaxation processes that are involved.

We are here specially concerned with the effect of $\langle\ \rangle$ on the basic POTENTIAL-FLUX equation

$$\begin{aligned} \text{FLUX} &= k \times \text{POTENTIAL DIFFERENCE} \\ &= k \times \text{P. D.} \end{aligned} \tag{1}$$

Obviously

$$\langle \text{FLUX} \rangle = \langle k \times \text{P.D.} \rangle \tag{2}$$

However $\langle \text{P.D.} \rangle$ is a fundamental quantity that is measured, so Eq. 2 is rewritten as

$$\langle \text{FLUX} \rangle = \langle k \rangle \times \langle \text{P.D.} \rangle + J \tag{2a}$$

where J is defined so that Eq. 2a is equivalent to Eq. 2.

Therefore

$$J = \langle k \times \text{P. D.} \rangle - \langle k \rangle \times \langle \text{P.D.} \rangle \tag{3}$$

$$= \langle \left[k - \langle k \rangle \right] \times \left[\text{P.D.} - \langle \text{P.D.} \rangle \right] \rangle \tag{3a}$$

If we note that $Q - \langle Q \rangle$ is a local dynamic fluctuation, and that $\langle PQ \rangle$ is a temporal covariance, we see that J measures the temporal covariance of local dynamical fluctuations in k and the P. D. When Eq. 2a is studied empirically, the term J is called active transport.

At macroscopic steady-state, $\langle \text{FLUX} \rangle$ is constant and can be taken as zero, so consequently

$$\langle \text{P.D.} \rangle = - J / \langle k \rangle \tag{4}$$

Thus active transport is able to create a non-equilibrium steady-state. Similarly at macroscopic equilibrium, $\langle \text{P.D.} \rangle$ is constant and can be taken as zero, so

$$\langle \text{FLUX} \rangle = J \tag{5}$$

Thus active transport is also able to create an equilibrium flux. Since the

concept of active transport is relatively new, and is usually considered only in the molecular domain, Eqs. 4 and 5 have not yet been considered in gross physiology. However these equations seem likely to figure prominently in the eventual understanding of many vital processes such as cytoplasmic streaming, the maintenance of non-equilibrium metabolic pools, etc.

Eq. 3a makes explicit the dynamical nature of active transport. Structural considerations are implicit in this formalism, becoming explicit only when specific models are proposed. For example, the activated carrier model of membrane transport posits a mobile carrier whose affinity for the transported species differs on the two sides of an intramembrane diffusion barrier. Obviously this model satisfies the dynamical requirement that the free energy fluctuations, associated with binding to the carrier, are temporally correlated with the conductance fluctuations, associated with a lowered diffusion barrier. However the model requires a variety of intramembrane structural specializations which may or may not actually exist.

For comparison we note that rectification,[1] as is produced at donor-acceptor junctions, provides an alternative to the carrier as a structural specialization which permits controlled conductance fluctuations. The active transport current of ions may well be generated by ATP-driven intramembranal voltage fluctuations which drive correlated conductance fluctuations in a neighboring rectifying junction. The chemical specificity of the rectifier acceptor sites would then replace the chemical specificity of the carrier.

Now that we have supplied a clarification of the dynamical nature of active transport at one level - the membrane level - it should not be too surprising to fina a similar result at higher levels, e.g. at the organ level. We might expect that the heart also generates active transport. For example, mammalian hearts use muscular contractions to generate pressure fluctuations which are correlated with the resistance fluctuations generated by the valves. The blood circulatory relaxation time constants are measured in seconds, rather than in milliseconds as they are for transmembrane ion transport. Yet the blood pump and the sodium pump both perform essentially the same dynamical task, as specified by Equation 3a.

The question now naturally arises: does there exist a dynamical (temporal) hierarchy of active transport systems which in some sense parallels the known structural hierarchy of pores, vesicles, microtubules, etc.? Indirect evidence fur such a dynamical hierarchy comes from the extensive self-documentation of driven fluctuations in a wide variety of physiological quantities covering a range of mean periods from milliseconds to months. So far, however, we have interpreted these fluctuations as serving time-dependent control and optimization function.

It might be suggested, more fundamentally, that whenever driven fluctuations occur in living systems, physiologically significant covariances as defined in Eq. 3a also occur. For example, red blood cells do not flow steadily

[1] See for example, Kornacker in (59).

in the microcirculation, but are switched from capillary to capillary with a mean period of some 10^2 seconds. Speculatively we might expect this to cause corresponding fluctuations in the local oxygen tension. Such fluctuations could contribute to an active transport of oxygen if there were correlated fluctuations in the oxygen permeability of the microcirculatory shrubs.[1] Also the oxygen tension fluctuations could contribute to the maintenance of non-equilibrium steady-state ATP concentrations if there were correlated fluctuations in the local rate constant for oxidative phosphorylation. Hopefully these brief suggestions will be of some use to experimental and theoretical physiologists working toward an understanding of vital processes.

[1] To be discussed later, the capillary shrubs are those capillary segments that are tied to the same arterial-venous point of origin and disappearance.

Bibliography

1. Sjostrand, F., Nature, 171, 30, 1953.
2. " , J. Cell. Comp. Physiol., 42, 15, 1953.
3. Danielli, J., Davson, H., J. Cell. Comp. Physiol., 5, 495 1935.
4. " , "Permeability of Nat. Membranes", Cambridge U. Press, 1952, London.
5. Robertson, J., J. Biophys, Biochem. Cytol., 3, 1043, 1957.
6. " , Biochem. Soc. Symp., 16, 3, 1959.
7. Elbers, P., Recent Prog. Surface Sci., 2, 443, 1964.
8. Mueller, P., Rudin, D., Tein, H., Westcott, W., Circ., 26, 1167, 1962.
9. Maddy, A., Huang, C., Thompson, T., Fed. Proc., 25, 933, 1966.
10. Bangham, A., Standish, M., Watkins, J., J. Mol. Biol., 13, 238, 1965.
11. Bangham, A., Standish, M., Weissman, G., J. Mol. Biol., 15, 253, 1965.
12. Rothfield, L., Finkelstein, A., Ann. Rev. Biochem., 37, 475, 1968.
13. Mitchell, P., Biol. Rev. Cambridge Phil. Soc., 41, 445, 1966.
14. Armstrong, W., Rothstein, A., J. Gen. Physiology, 48, 61, 1964.
15. Maddy, A., Malcolm, B., Sci., 150, 1616, 1965.
16. " , Sci., 153, 213, 1966.
17. Lehninger, A., Neurosci. Res. Bull., 6, 15, 1968.
18. Cook, G., Heard, D., Serman, G., Nature, 191, 44, 1961.
19. Cook, G., Nature, 195, 159, 1962.
20. Eylar, E., Madoff, M., Boody, O., Oncley, J., J. Biol. Chem., 237, 1992, 1962.
21. Hodgkin, A., Keynes, R., J. Physiol., 128, 28, 1955.
22. Caldwell, P., Hodgkin, A., Keynes, R., Shaw, T., J. Physiol., 152, 561, 1960.
23. Whittam, R., Biochem. J., 84, 110, 1962.
24. Weed, R., LaCelle, P., Merrill, E., J. Clin. Invest., 48, 795, 1969.
25. Bensch, R., Bensch, R., Nature, 221, 618, 1969.
26. Kundig, W., Ghosh, S., Roseman, S., Proc. Nat. Acad. Sci., 52, 1067, 1964.
27. Francoeur, M., Denstedt, O., Canad, J. Biochem. Physiol., 32, 664, 1954.
28. Keilin, D., Biochem. J., 49, 88, 1951.
29. Simoni, R., Leventhal, M., Kundig, F., et al, Proc. Nat. Acad. Sci., 58, 1963, 1967.
30. Robertson, S., in Locke M.(ed)."Twenty Second Symp. of the Soc. Study Dev. and Growth", Academic Press, N.Y., 1964, pp. 1-81.
31. " , in Wolstenholme, W.(ed.)"Ciba. Symp. Unit Membrane", Little Brown Co., Boston.

32. Schmitt, F., Exp. Cell. Res., Suppl. 5, 33, 1958.

33. Sjostrand, F., J. Ultrastructure, 9, 340, 1963.

34. " , J. Ultrastructure, 9, 561, 1963.

35. " , Nature, 199, 1962, 1963.

36. " , Protoplasma, 63, 248, 1967.

37. Wallach, D., Gordon, A., Feder. Proc. 27:1264, 1968.

38. Lenard, J., Singer, S., Pro. Nat. Acad. Sci., U.S., 56:1828, 1966.

39. Karn, E., Feder. Proc. 28:6, 1969.

40. Green, D., Perdue, J., Ann. N.Y.Acad. Sci., 137, 667, 1966.

41. Fernandez-Moran, H., Circ., 26, 1039, 1962.

42. Fernandez-Moran, H., Oda, T., Blair, P., Green, D., J. Cell. Biol., 22, 63, 1964.

43. Passow, H., in Bishop C., Surgenor D., (eds.)"The RBC", Academic Press, N.Y., 1964, Chapt. 3.

44. Keynes, R., Quart. Rev. Biophysics, 2, 177, 1969.

45. Schultz, S., Curran, P., in Handbook of Physiol.III, Amer. Physiol. Soc., Wash., 1968; Section 6, 1245-1275.

46. Wilbrandt, W., Rosenberg, T., Pharmacol. Rev., 13, 109, 1961.

47. Ling, G., "Physical Theory of the Living State", Blaisdell, Waltham, Mass. 1962.

48. Troshim, A., in Kleinzoller, Kotyk, A., "Membrane Transport and Metabolsim", Academic Press, N.Y., 1961, pp. 45-53.

49. Harris, E., Pankerd, T., J. Gen. Physiol., 41, 197, 1957-58.

50. Miller, D., Biochem. Biophys. Acta, 37, 448, 1960.

51. Bratton, C., Hopkins, A., Weinberg, J., Science, 147, 738, 1965.

52. Nachmansohn, D., Proteins in Excitable Membranes, Science, 168, No. 3935, 1059, 1970.

53. Davson, H., "A Textbook of General Physiology", Little, Brown and Co., Boston, 1964.

54. Dick, D., "Cell Water", Butterworths, 1966.

55. Loewenstein, W., (ed.) Biological Membranes: Recent Progress, Ann. N. Y. Acad. Sci., 137, 403-1048, 1966.

56. "Membranes and Transport Phenomena", 10th Annual Meeting Biophysical Society, Feb. 1966.

57. Stein, W., "The Movement of Molecules across Cell Membranes", Academic, 1967.

58. Jörnefelt, J., "Regulatory Functions of Biological Membranes", Elsevier, 1968.

59. Dowben, R., (ed.) "Biological Membranes", Little, Brown and Co., Boston, 1969.

60. Zweifach, B., Microvas. Res. 1, 83, 1968; 2, 96, 1970.

61. Bruns, R., Palade, G., Studies on Blood Capillaries , Parts I and II, J. Cell Bio., 37, No. 2, 244-276, 277-299, 1968.

62. Loewenstein, W., Cell to Cell Passage of Large Molecules, Nature, 212, 629, 1966; Intercellular Communication, Sci. Amer., 222, 79, 1970.

63. Sha'afi, R., Rich, G., Mikulecky, D., Solomon, A., J. Gen. Physiol., 55: 427, 1970.

VI. SOME COMMENTS ON CONTROL OF THE MICROVASCULAR SYSTEM

If a system's view is murky at the membrane level, form and function emerge with more clarity at the next higher organizational level in the complex biological system, the microvascular system. Its lateral dimensions range from membrane magnitude (a fractional micron endothelial cell thickness constituting the wall) to cellular extent (5-10-20-30 micra diameters). However folded up, it makes up much of the body of the organs (e.g., liver, kidney, lung); and in length and extent it spans the body.

1. Introduction

The purpose of the following analysis is to examine some aspects of the role of the microvascular system in maintaining homeokinesis of the whole animal, i.e., in man and other vertebrates. Particular emphasis will be placed on (a) the dynamics of form (the morphokinesis of the system), (b) the mechanisms that have been suggested for controlling the system (the dynamics of function) with their attendant successes and problems, and (c) some recommendations for experiments that may further elucidate the control mechanisms.

2. Background

The circulatory system may be considered to have two main functions, both of which can be described as transport and communications functions. Both functions are required if homeokinesis is to occur and be maintained. These two transits occur via the permeability characteristics of microvessels and via the movement or circulation of the blood. The major permeability or transvascular transit occurs via capillaries (or sinusoids) and immediate post capillary venules and the communication transit between organs is accomplished by the circulating blood. These transit functions are closely coupled to the central and peripheral parts of the nervous system, the metabolism of organs and the functions of the lymphatic system. As yet only the broadest outlines of these transit functions have been delineated in respect to the cardiovascular system. Furthermore the bulk of the quantitative data that has been secured for control mechanisms has been from the gross vascular system; the data from the microvascular system are scanty and essentially qualitative and species limited. Too often there is poor correlation between the control mechanisms that have been experimentally derived from the gross and from the microvascular systems. As yet there are no published data which reconcile and describe the control mechanisms in both systems simultaneously; preferably as secured in an unanesthetized animal. (The excellent reviews of the control of the 'peripheral' circulation that have been published during the last decade have been written by gross physiologists who have utilized, wholly or to a large extent, data that were secured by gross physiological methods and 'fitted' to this information some rather divergent data that were secured by methods of direct microscopy of small blood vessels. (Green and

Kepchar, (1); Folkow, (2); Gauer and Henry, (3); Guyton and Coleman, (4); Haddy and Scott, (5); Mellander and Johanssen, (6).) Thus, the conclusions pertaining to the operational characteristics of the microvascular system have been 'deduced'. While many of the deductions may prove to be valid, there is too much of the unknown or long chain inference present. The only possible exception to this assertion is a description of the relation between topology and geometry of the microvascular and gross arterial vasculature at sizes above capillaries[1].).

Before the control of small blood vessels will be discussed the microvascular system will be defined and some common attributes of these vessels will be considered.

All small blood vessels which require microscopy for their recognition are designated as belonging to the microvascular system; this term includes the blood flowing in them. The term 'microcirculation' is restricted to the blood in the microscopic vessels.

The upper limit of the diameter of blood vessels comprising the microvascular system in vertebrates is 500 micra (Bloch (7)). This number was derived from the visual resolution of the normal human eye, which is approximately 100 micra, and the use of the concept of the signal to noise ratio concept, chosen at the level of 1:5. That is, an object, in this case the smallest vessel, that can be recognized by the human observer measures approximately 100 micra, but the nature of the vessel cannot be determined. It will have to be enlarged approximately five times before unequivocal identification is possible. In some vertebrates, as in a rodent, a vessel with a diameter of 500 micra would be a major vessel, while in the Congo eel (Amphiuma means) the vessel would belong to 'microscopic' vessels as it would be about six diameters wide in respect to the largest dimension of that animal's red-cell, which measures approximately 80x25 micra. Thus the microvascular system, as it is defined, consists of many different types of vessels of which the vessels which are concerned with transvascular transit are but part of the system.

The above classification should be related to and compared with a more general classification of the vessels of the cardiovascular system which has been derived from gross physiological studies. One such classification (Heymans and Folkow, (8)) divides the system into " 'Windkessel' vessels (e.g. aorta) which convert a rhythmic input into a fairly smooth outflow. Resistance vessels, each of which contains an arteriole, capillary and venule, the sum of which result in the total resistance to flow and their interrelation results in one of the major determinants of transvascular transit. Exchange vessels whereby blood comes in 'contact' with tissue via an enormous

[1] This description has been elucidated through the work of Mall (1888-1906); Green (1944); Groat (1948); Suwa (1963); Iberall (1967). In an 'average' sense, the character of branching from the aorta down to tubes of the order of 20 micra are known throughout the mammalian arterial circulation. See addendum to this section.

capillary surface. Sphincter vessels, occurring in the terminal arterioles, where smooth muscle 'sphincters' can induce intermittent occlusion thus affecting the capillary surface available for perfusion. Capacitance vessels, where shifts in average bore, so moderate that their flow resistance is but little affected, but which can profoundly affect volume flow and cardiac filling, and, shunt vessels, the arterio-venous anastomoses".

In accord with our proposed systems' views, more grossly we may also regard the circulatory system as divided into a source and supply system, and a return system. These systems are functionally concerned with

Source and Supply System

Generation of pressure and flow
'Long line' transmission
Functional distribution
Local exchange and supply

Return System

Functional collection
Large volume storage

The microvasculature is concerned with local supply and exchange.

It is meaningless when control is being considered to consider vessels per se, as vessels do not exist by themselves but are imbedded in unique environments which differ from each other, and therefore their environments must also be considered. This means that paradoxically there is no 'microvascular system', when considered as a functional entity, but that there are 'microvascular systems'. The paradox is resolved by discussing microscopic blood vessels as components of functional units of organs.

The bulk of what makes up the microvascular systems consists of arterioles, capillaries and venules. In addition to these vessels two kinds of shunts exist between the arteriolar and venular systems, namely, "preferential" or "thoroughfare channels" and arterio-venous anastomoses. The "preferential" channels are vessels that function as capillaries, in regard to transvascular transit, but differ from "true" capillaries by having scattered smooth muscle-like cells encircling them which can alter their lumen. In general, the preferential channels have blood flowing through them continuously while the capillaries, derivative from these channels, can be and are frequently excluded from active flow (Zweifach and Metz (9)). The other type of shunt is the arterio-venous anastomosis, which in contrast to preferential channels, is short in length, muscular, and opens only intermittently. Most of the time these vessels are closed. In respect to the spatial localization of arterio-venous anatomoses they are found in the liver, gastro-intestinal tract, skeletal muscle[1] and connective tissue but not in the nephron. Prefer-

[1] There is some question about the extent and location of arterio-venous anastomoses in skeletal muscle. They have been described in the anatomical literature, although the literature is sparse (Clara (10)).

ential channels exist in connective tissue (mesenteries of the rodent where they have been studied extensively), striated muscle, gastro-intestinal tract, but do not occur in the liver or nephron.

The longitudinal configuration of microvessels differ. Arterioles are cone-shaped vessels whose taper is approximately 1° with the direction of the smallest dimension of the taper toward the capillary bed, while venules have a similar taper but their diameter increases as they leave the capillary bed.[1] Capillaries and preferential channels are essentially cylindrical. The taper of arterioles may add to the frictional resistance of the cellular components of the blood, but this factor has not been measured. The taper may have some importance in those human diseases where erythrocyte aggregation occurs, as the aggregates must, and usually do, deform as they are forced toward the capillaries; but in so doing, their linear velocity progressively decreases (Bloch, 1962; Jeffords and Knisely, 1956).

Microscopic vessels also exhibit dynamic fluctuations in their internal diameters which may be especially notable in arterioles. This phenomenon has been described as "vasomotion" but will not be used in the discussion because the term has been used to describe two different responses, and most investigators use the term indiscriminately so that it is frequently impossible to know what the writer has in mind. For example, Zweifach (1955), who has studied the phenomenon primarily in rodents describes vasomotion as "the spontaneous opening and closure of the precapillary sphincter. The term 'vasomotion' has been applied to this phenomenon to set it apart from the wave-like patterns of constriction and dilation exhibited by the larger blood vessels." ((9), p. 283, 1955). On the other hand Nicoll and Webb (13) define "vasomotion", as observed in the living membranes of bats "as changes in the lumen diameter that result from activity of the muscle cells in the vessel wall. Arteries and occasionally the largest arterioles exhibit active vasomotion in two distinct ways. One is a slowly developing caliber change ... (in contrast to) active vasomotion (which is) a rapid caliber change." Rapid vasomotor changes occur as peristaltic waves that "move from arcuate arterioles toward terminal arterioles: while localized contractions are exhibited most characteristically by the 'final cells' of terminal arterioles, the 'precapillary sphincters'"((13), p. 295, 1955).

The impressions that are commonly held in gross physiology with regard to this phenomenon may be gleaned from the following well-known gross physiology texts.

"The exchange between blood and tissues takes place in the true capillaries, through which both the amount and direction of blood flow are adjusted by variation in the caliber at their points of origin from the thoroughfare channel. The opening into each true capillary is guarded by a cuff of smooth muscle (the precapillary sphincter) which tends to regulate the flow of blood into the capillary it serves. At intervals of 30 seconds

[1] In general, vessels in the arterial tree taper not more than one to two degrees as they descend from the aorta down to arterioles. (Jeffords and Knisely (11); Suwa (12.)),

to 2 minutes, some of the precapillary sphincters open up while others close down to produce a continuously changing pattern of blood flow, rate and direction through the individual capillary segments. The phasic diameter changes of controlled resistance are termed vasomotion ... The major control over regional blood flow distribution is exerted by control of the caliber of terminal arteries, arterioles and precapillary sphincters." (Rushmer in Ruch and Patton (1966) p. 605.)

"Blood does not flow at a continuous rate through the capillaries. Instead it flows in intermittent spurts. The cause of this intermittency is the phenomenon called vasomotion, which means intermittent contraction of the meta-arterioles and precapillary sphincters. These constrict and relax in an alternating cycle 6 to 12 times per minute." (Guyton (1966) p. 434.)

"The functional anatomy of the capillaries has been worked out for some cold-blodded animals, for the rat and dog mesentery ... and for the myocardium. Briefly, blood flows ... from the arterioles ... into a metarteriole, and then into capillaries. ... At the ostia of each capillary is a small precapillary sphincter of smooth muscle which is controlled by nerves presumably from the sympathetic nervous system, in the same manner that these nerves control the arterioles and metarterioles. In the body, the metarterioles undergo periodic contractions at intervals of 15 seconds to 3 minutes. When the tissue is in a resting state, the constrictor phase of this rhythm predominates and the precapillary sphincters may be completely closed. When the tissue becomes active, the dilator phase of the metarterioles predominates and the precapillary sphincters are open. ..." (Gregg in Best and Taylor (1961) p. 158.)

However, whether direct observation of precapillary sphincter action is actually observed in muscle is moot. For example, Honig et al.(14) undertook to investigate whether tonic neural discharge might contribute to the control of capillary density by counting the density of open capillaries in resting rat gracilis muscle before and after acute denervation. It appeared that vasoconstrictor tone influenced arterioles but not precapillary sphincters. "Denervation produced no change whatever in the number of open capillaries, whether one field was observed continuously for 10 min. after denervation ... or data for all fields were pooled ..." They make the point that precapillary sphincters are innervated by autonomic nerves, based on visualization of myoneural junctions observed by electron microscopy. Yet, operationally, their evidence for precapillary sphincter action in muscle was inferential. "In every rat the precapillary sphincters exhibited contractile activity, measured by spontaneous, random changes in capillary density ..."

Further, even though "normal vasomotion (opening and closing of capillaries)" was at issue, the criterion was not direct observation of the changing diameter of either of the capillary entrances or the capillary lumen, but whether they "contained moving erythrocytes at that instant" of observation. If for example, the capillaries contained only plasma, as a result of variable skimming upstream, an inference of precapillary vasomotion may be unwarranted.

On the other hand, a study of capillary flow in muscle reported by Cardon et al. (15) indicates a periodic character to the number of red cells passing through small capillaries in the panniculus muscle of the mouse. Similar results of earlier studies on capillaries in the frog mesentery and the rectus femoris muscle in the guinea pig are mentioned. They conclude "A morphological source for the oscillations in red cell flow was not established ... at no time were changes noted in the intraluminal diameter of arterioles or capillaries, or of sphincteric action sufficient to interfere with red cell flow..."

In fact, there is no morphological evidence that fluctuations in precapillary inlet diameters is the common (or a very common) means for controlling capillary flow. On the other hand, fluctuations in red cell flow in small capillaries (capillaries smaller than red cell diameters) seems to be quite general.

Maintained changes in the diameter of blood vessels for varying periods of time are common to all living vessels and this response is termed "vascular tone". Maximum tone exists when the vessel is constricted to zero lumen, while tone is absent when the vessel attains its maximum diameter. In maintaining "tone" the basic longitudinal configuration of that particular type of vessel is also maintained which is characteristic for arterioles, capillaries and venules. The manner in which vascular tone is produced is being elucidated. Tone may be considered as the result of integrated phasic twitches initiated by action potentials. (Funaki (16); Steedman (17); Siggins (18).) This electrical activity is apparently myogenic, at least in some vessels; it is not dependent on the nervous system (Johansson and Ljung (19)). However, other vascular smooth muscle do not exhibit automaticity, i.e., synchronized contractions (Mellander and Johansson (6)). Furthermore, the contractile elements of still other vascular smooth muscle do not exhibit action potentials but produce their activity through graded changes in membrane potential (Su, et al. (20)). Thus vascular tone may be the product of different mechanisms in various regions of the vascular system as: (a) Myogenic automaticity and propagation which result in rhythmicity, modulated by vasomotor nerves, vasoactive substances or local metabolites; (b) Tone may be developed by asynchronous spike discharges, the result of vasomotor nerves; and (c) Tone may be produced by changes in membrane potentials producing graded contractures. All these mechanisms might operate in one vessel, each to a varying degree. (Mellander and Johansson (6).)

Other common attributes of small blood are their morphological components and structural organization. All blood vessels, including capillaries, have three 'tunics', that is, an intima whose major component consists of endothelium, a media which is composed of smooth muscle or 'modified' smooth muscle cells and an adventitia which consists of connective tissue (Rhodin (21)). The proportions of these tunics vary: capillaries and sinusoids are mostly constructed of endothelium, their media consists of a few scattered cells (pericytes or 'modified' smooth muscle cells) and they have very little adventitia, a few 'strands' of connective tissue. A prominent feature of many arterioles is their media which consists of smooth muscle. (Electron microscopy is delineating structural details of the media which may have implications in regard to a receptor mechanism for humoral transmitters. For

example, arterioles with a post vital diameter of 30 micra or less (rabbit, fascia of the thigh) have smooth muscle cells and endothelial cells that make membrane to membrane contacts, that is, the basement membranes are absent. Rhodin (21) suggests that these intimate contacts serve as channels that expedite the transfer of vasoactive substances from the blood to the effector site in smooth muscle cells. Thus these endothelial cells can be considered as receptors that initiate the depolarization of smooth muscle cells.) Arterioles in general have more smooth muscle than comparable sized venules. There are exceptions. A notable exception are the terminal portions of the hepatic arterioles which lack smooth muscle cells but are contractile (e.g., McCuskey (22)). Also, as a general principle, the vessels of the venous system show considerable variation in smooth muscle, both in amount and location. In regard to the latter smooth muscle may occur in the adventitia.

An important structural constituent of vessels, especially in regard to any control mechanism, are their nerves. Unfortunately the extent of the distribution of nerves in the vascular system is still not known adequately. While the anatomical distribution of nerves to blood vessels has been studied for over a century, the methods, until the last decade, have delineated not only nerves but connective tissue fibers. Furthermore, it has not always been possible to distinguish, morphologically between sensory and motor fibers. Hence, the mapping of the nerve supply to blood vessels is far from complete and is, to some measure, an open question. Too little data exist which unequivocally demonstrate how peripheral nerves terminate. What is being delineated is the manner of the relationship of the nerves with the microvascular system. The best information in regard to the sympathetic nerves has been secured in the past decade with electron and fluorescent microscopy.[1] While the smallest nerve fibers, neurites, cannot be traced or separated from connective tissue fibers with certainty even with these methods, the location of neurotransmitter substance (norepinephrine) has been identified along the course of sympathetic nerves. Sympathetic nerves penetrate the adventitia and reach the outer border of the media, but apparently do not enter it, or do so only to a very limited degree (see below) and in larger vessels of the microvascular system there is no 'intimate' contact of vesicles with smooth muscle cells. The vesicles are separated from smooth muscle cells by distances of hundreds or thousands of Angstrom units. (In terminal arterioles, unmyelinated nerve fibers are particularly numerous in the region of precapillary sphincters. Some of the axons contain vesicles with densely staining granules (norepinephrine). Furthermore, occasionally 'club' like nerve endings make close contact with vascular smooth muscle (Rhodin (21)). However, it is still difficult to unequivocally separate sensory from motor nerves at this level of the microvascular system.) Thus in microvessels which have multiple layers of smooth muscle enveloping the endothelium, the innermost layers would be exposed to the lowest concentration of a neuro effector substance. There is less information about the topographic distribution of parasympathetic nerve fibers with microvessels. The current opinion relegates the parasympathetic to the cranial and caudal ends of the neuraxis

[1] The following section presents an up-to-date but also speculative view of the autonomic nervous system.

and there continues to remain a vexatious problem, the possible existence of efferent vasodilator 'fibers' (parasympathetic) that supposedly emerge in dorsal spinal nerve roots (dorsal root efferents). If these efferent vasodilator fibers do exist then the parasympathetic system is far more widespread, and the generally current concept is incorrect.

Finally the vessels appear to contain or be associated with sensory nerves. Again detailed information is lacking in respect to their distribution with microvessels, let alone what their physiological role might be. It is doubtful that they are sensors only of pain and pressure; perhaps they are chemical sensors.

In summary, morphological information about the relation of the nervous system with the microvascular system is incomplete. Probably the most extensive information exists for the extremities, iris, gastrointestinal tract and heart; but these structures have been studied insufficiently or not at all, with dynamic methods, in regard to the controlling microvascular systems.

In spite of the too many anatomical and physiological problems that remain for the interrelationships of the nervous and vascular systems, the influence of the nervous system on the cardiovascular system is profound. From gross physiologic evidence, the sympathetic system is probably the most significant system that influences the cardiovascular system, and the gross physiological studies have resulted in various classifications of the autonomic nervous system such as the one cited. (A classification of autonomic nervous system as affecting the cardiovascular system is proposed by Heymans and Folkow (8), "Centrally Controlled Vasomotor Fiber Systems. Sympathetic Adrenergic Vasoconstrictor Fibers: These are the only vasomotor fibers that are tonically active; they participate in practically all types of neurogenic cardiovascular adjustments. They form the sole efferent motor pathway for the control of the level of the blood pressure and for the maintainance of the temperature equilibrium. They also constitute the basis for the most common type of neurogenic vasodilation, i.e., simply by inducing inhibition of their tonic activity. Sympathetic Cholinergic Vasodilator Fibers: These are distributed to the skeletal muscles and possibly to the myocardium. They possess no tonic activity, but do engage in a specific somatomotor-visceromotor reaction pattern, probably elicited in emergency situations whereby the cardiovascular system is rapidly adjusted to increased blood flow to the muscles. (Adrenal Medulla-these hormones 'lend support' to the nervous pathways: Their excitatory effects are far inferior to those of the vasoconstrictor fibers: their inhibitory effects are probably more potent as 'even' small concentrations of epinephrine may exert a vasodilator effect, in some tissues, especially skeletal muscle.) Parasympathetic Cholinergic Vasodilator Fibers: These are distributed only to a few hemodynamically insignificant areas, in which they increase the blood flow. Peripherally Controlled Vasomotor Fiber Systems. Dorsal Root Vasodilator Fibers. Present in skin and superficial mucous membranes, where they create a local increase of blood flow in response to harmful stimuli; they possess no motor function whatever. They are identical with unmyelinated C fibers, are primarily concerned with pain, and involved in the axon reflex mechanism. Possible Local Nerve Cell Plexuses in the Vascular Walls. A Hypothetical System for Local Integration."

In summary the opinion has been expressed that "the tonic vascular control in circulatory homeostasis is executed only by the constrictor fibers, while the different types of dilator fibers, which do not form a homogeneous group, are reserved for a few, highly specialized purposes." ((8), p. 428.)

In an attempt to clarify the responses of vascular smooth muscle to adrenergic and cholinergic drugs hypothetical morphological receptor sites in vascular smooth muscle have been proposed (Ahlquist, (23), (24); Green, (25)). The alpha (constrictor) and beta (dilator) receptors are affected by adrenergic drugs, but only the alpha receptors are affected by stimulating sympathetic nerves. The alpha receptors appear to be widely distributed throughout the body and are most responsive, in decreasing order, in vessels of the kidney, skin, skeletal muscle, heart, spleen and liver. The vessels of the brain may be lacking in alpha receptors. Beta receptors are extensively 'distributed' in skeletal muscle, heart and spleen but sparse in the kidney and liver. (The latter would be contradictory to morphological evidence as there is no convincing evidence that dorsal root efferents exist. If they do exist, then they would belong to the parasympathetic system (vide supra, page 9) (Mitchell, (25a)).) Finally the receptors appear to be integrated through the sympathetic nervous system emanating from the medullary portion of the neuraxis.

3. The Organizational Level of Microvascular Control

Gross physiological studies of the control of peripheral resistance for blood flow indicate differences for different organs (e.g., Green (1)). The source for these differences probably resides in the smallest volume of tissue that makes that organ unique. The maintenance of their unique functional properties depends not only on receiving an 'adequate' volume of blood per unit time but equally important in regulating this volume of blood. The maintenance of homeokinesis of the whole organism must begin with the unique structure that constitutes each particular organ. Then the various stimuli that arise in an organ affect all others and these effects are communicated via the circulation and nervous system. The volume of particular tissue which engenders these responses is the <u>functional unit</u>.

The morphological components of the functional unit must contain all cellular components which make that organ unique. The components of the unit are analogous to the atoms of a molecule which make a molecule unique, e.g., the unique functional unit of sugar and salt. There are just so many components that cooperate to maintain a functional unit; redundancy does not occur. The redundancy which exists is in the number of functional units that make up an organ.

As yet the functional units of all organs have not been defined. To date the best definitions of a functional unit exist for the kidney - the nephron - and the nervous system - the neuron. In regard to the latter, the usual definition is incomplete, to wit, "the complete nerve cell, including the cell body, axon and dendrites" (Gould, (26)) because the vessels which are associated with the unit have been excluded. The exclusion of the micro-

vessels would make the unit nonoperative. Another illustration is the liver where the problem of defining the functional unit has existed for many years. For at least half a century it has been recognized that the lobule is not the smallest mass of tissue that makes this organ unique (e.g.,Mall (27) 1906). Recently the unit has been defined as follows: Its core consists of the sinusoid with its afferent connections to an interlobular portal venule and hepatic arteriole, and, an efferent connection to the central or sublobular hepatic venule. In addition, the unit contains the perisinusoidal space and connecting lymphatic(s), bile cannaliculi, sensory and autonomic nerves and a mass of hepatic tissue. The center of this core of tissue is circumscribed by a radius that originates in the center of a sinusoid and extends to the center of the immediately adjacent hepatic cells. In frogs the radius extends through the whole cell as two hepatic cells exist which separate adjacent sinusoids rather than one as in mammals (Bloch (28)). Furthermore, the length of the unit can and does alter due to local, regional, or external (systemic) demands without compromising the functional unit as it is defined (see Fig. 1, Bloch (7), (28)).

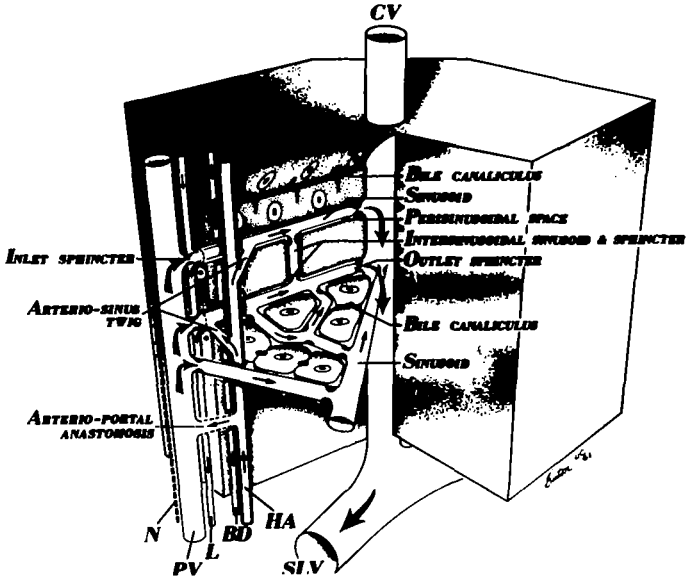


Fig. 1. The Hepatic Lobule (Bloch (7))

N	Nerves (Sensory, Sympathetic and Parasympathetic)
PV	Portal Venule
L	Lymphatics
BD	Bile Ductule
HA	Hepatic Arteriole
SLV	Sublobular Venule (Sinusoids also empty into this vessel.)

Another illustration is the functional unit of the lung. It is not the alveolus. The core of the unit is an alveolar capillary with its connections to the pulmonary arteriole and venule, the associated nerves, and adjacent epithelia. The volume of tissue consists of a radius that extends from the center of the alveolar capillary to one-half the distance to an adjacent parallel capillary.

The operational characteristics of the hepatic functional unit have been studied with the microscope perhaps more extensively than any other functional unit. It has been studied extensively in the living frog (Knisely, Bloch and Warner (29)). Selected aspects of the living unit have also been studied in the guinea pig (Irwin and McDonald (30)) and rat (McCuskey (22); Debaker (31); also see Bloch (32)). It was found that each functional unit can operate independently or in conjunction with any number of such units. Blood flow to the unit is controlled by inlet sphincters that are located at the junction of a sinusoid with the interlobular portal venule (Figs. 1 and 2). In addition, blood is supplied to a sinusoid from arterioles which join a sinusoid distal to the inlet or afferent sphincter. Both the inlet sphincter and the arteriole can modify their orifices independently. Outflow from a sinusoid can occur by one of two routes. In one, the blood in the sinusoid enters a central venule directly, that is the flow pattern is directly via one sinusoid from the interlobular portal venule through a sinusoid and into a central venule. An alternative pathway is via an intersinusoidal sinusoid, that is, a sinusoid that interconnects sinusoids. Here the flow pattern is from a portal venule via a sinusoid that joins the portal venule, which then joins an intersinusoidal sinusoid, to connect with a sublobular hepatic venule (Fig. 2E). In either case, the flow of blood into the central or sublobular venule is controlled by an outlet (efferent) sphincter which can act independently of the inlet or intersinusoidal sphincters. These sphincters can hold the blood in a sinusoid for varying periods of time, and when they release the blood into the outflow tract, the inlet sphincter can close and the sinusoid 'milks' the blood into the hepatic venous system. Thus each sinusoid acts as a minute blood reservoir. The response has been given the name of "autotransfusion" (29). In addition to the above reaction, the wall of a sinusoid has several degrees of permeability; at one extreme most of the plasma may be removed so that the blood consists of packed cells, the other extreme is when there is no apparent change in the cellular concentration of the stored blood, that is, the cell concentration is identical with that in a portal venule. The wall of the sinusoid, including the sphincters, can also participate in another reaction, the phagocytosis of particulates, seemingly without interfering with their other functions. Thus the hepatic functional unit illustrates local control as evidenced by phagocytosis and permeability, and regional and systemic control by working together with associated units as exhibited by the autotransfusion reaction. The unit controls linear and volume flow, permeability, lymph production, bile formation and the multiple functions of the hepatic cells, thereby displaying unit and systemic integrated control.

The description of the hepatic functional unit contains two concepts that require further examination, namely, the relationship of the metabolic activity of an organ and the number of capillaries with blood flowing through them and the concept of the independent contractility of capillaries.

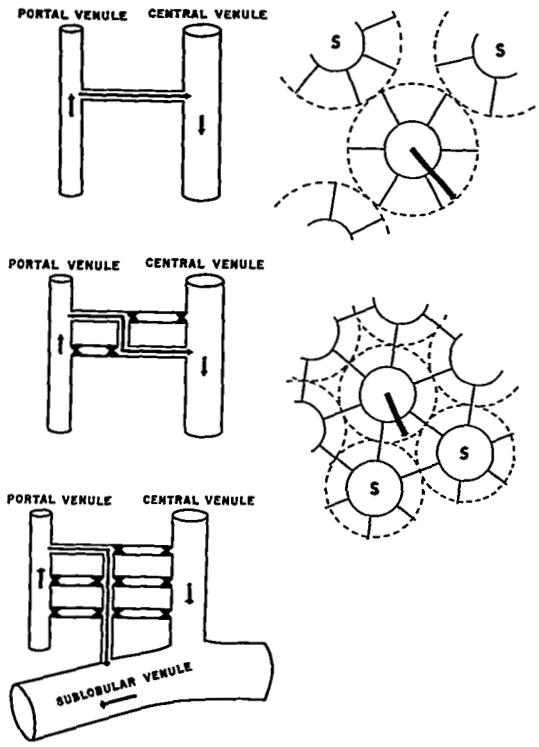


Fig. 2. Blood Flow Patterns and Maximum Diffusion Distance of the Hepatic Functional Unit (Bloch (7), (28))

A. The blood flow pattern in a sinusoid connecting a portal and central venule.
B. The maximum diffusion distance, indicated by the arrow from the center of a sinusoid (in amphibia) to the periphery of a hepatic cell (HC).
C. The blood flow pattern from portal to central venule passing through an intersinusoidal sinusoid.
D. The maximum diffusion distance, indicated by the arrow, from the center of a sinusoid (in the mammal) to the center of a hepatic cell (HC).
E. The blood flow pattern from a portal venule through a sinusoid, then through two intersinusoidal sinusoids, then into another sinusoid and emptying into a sublobular venule.

The relationship between the number of vessels with blood flow per unit volume of tissue and metabolic activity has been examined extensively in skeletal muscle. For example, the volume flow in skeletal muscle of the human forearm at rest may increase ten fold following exercise. This hyperemia of work is considered to be due to vasodilatation of previously closed vessels as well as dilatation of vessels that were open prior to exercise (e.g., Barcroft and Swan (33)). Such data have been secured by gross physiological methods (e.g., plethysmography) and from post-vital examination of skeletal muscle in experimental animals after the injection of markers (Krogh (34)). In respect to the latter, Krogh compared the number of capillaries per square mm. of muscle at rest and immediately after or during activity, or with muscles that exhibited little activity with those that were almost continuously active. For example, he reported that the extensor tarsi muscle (frog) at rest had an average capillary count of $5/mm^2$ compared to $195/mm^2$ during activity, while the rectus abdominus muscle of the guinea pig had a count of 200 capillaries/mm^2 compared to 2500 capillaries/mm^2 for the diaphragm. Furthermore Krogh found that the average diameters of capillaries in working muscle were greater than at rest, 4.3 micra vs 6.8 micra (frog) and 3.5 micra vs 5.0 micra (guinea pig). These examples suggest the presence of some metabolite(s), the excessive utilization of supply of some nutrient, or the nervous system affecting these vessels.

The vasomotor control of skeletal muscle in man arises from various sources. For example, the abolishment of vasoconstrictor nerves to human skeletal muscle doubles blood flow (Barcroft (33); Shepherd (35)), while the activity of vasodilator fibers as indicated by increased blood flow through the muscles following emotional stress and flow is also modified by posture, presumably via receptors in the thorax (Shepherd (35)). The extent of the influence of vasomotor nerves on blood vessels in other tissues have not been assessed as extensively in man as those in the limbs. While there is considerable influence of the nervous system on the vessels in functional units of organs, the data that illustrates such differences as well as the quality of the data varies.

There is a large body of data for 'vasoactive' substances which have been demonstrated during the last century (see the review of Haddy and Scott (5)). The sum of these studies indicate many promising candidates, but there is apparently no one substance that explains all of the observed effects. In all probability, it is a combination of substances, and these probably vary in importance in different organs. The most convincing candidates for influencing vasomotor activity, directly or indirectly, as established from gross physiological experiments, are: oxygen, changes in pH, potassium, adenine compounds, acetylcholine, lactic acid, bradykinin, histamine, and catecholamines.

Few modern studies exist where attempts have been made to elucidate the control of small vessels by direct microscopy. The classical studies of variations in blood flow through vessels in situ with the microscope were made by Sir Thomas Lewis (36) in human skin vessels, who considered the major vasodilator to be a histamine-like substance, and Ebbecke (37a), who deduced the responses of minute vessels in organs from gross reactions; Krogh (34),

who examined the dynamics of the circulation in frogs (tongue and membranes) and considered pituitrin the important substance for controlling permeability; Richards and associates (37), who examined the frog kidney (1922-1939), and concluded that epinephrine was important for controlling flow through glomeruli; and Fulton and Lutz (38) who, examiningthe vessels in the retrolingual membrane of the frog,elucidated the vasodilator role of nerves in affecting living microscopic vessels. More recently the effect of glycogenolytic substances and adenine compounds as well as oxygen have been examined as they affect the hepatic microvascular system in living frogs and rats using the method of in vivo microscopy from which a theory of vascular control was derived (McCuskey (22)). The study confirmed previous investigations where it was found that the sphincters of the hepatic unit were quite responsive to vasoactive substances (Bloch (39)). It was found that a correlation existed between the effectiveness of substances that stimulated glycogenolysis and their effectiveness in dilating vessels of the hepatic unit. (Oxygen concentrations, above normal, were found to constrict hepatic arterioles.) Furthermore, adenosine and adenine nucleotides, intermediates of glycogenolysis, were also effective dilators of the hepatic microvascular system. These results, derived from direct microscopy of living livers, in frogs and rats, coupled with current biochemical data resulted in a theory for regulating blood flow through the hepatic functional unit which is presented schematically in Fig. 3. The theory suggests that glucagon, isopropylnorepinephrine, epinephrine and norepinephrine stimulate an increase in the enzyme adenyl cyclase which converts 5'-ATP to cyclic -3', 5'-AMP. The subsequent accumulation of the latter activates a kinase to convert inactive phosphorylase to active phosphorylase resulting in the degradation of glycogen to glucose in the hepatic cell. The glucose that is released into sinusoids, while increasing the concentration of this substance by several hundred per cent apparently has no effect on vessel diameters. The cyclic-3', 5'-AMP that is produced by glycogenolysis does not accumulate in large amounts as it is degraded to 5'-AMP by phosphodiesterase, and adenosine is produced rather than inosine due to the presence of adenylic deaminase. (Adenosine and adenine nucleotides are potent vasodilators of hepatic microvessels.) Concomitant with glycogenolysis is a large efflux of potassium prior to the elevation of the blood sugar. The efflux of potassium doubles and perhaps triples the normal plasma level in hepatic venous blood and this ion,while a potent dilator of hepatic vessels, does not affect hepatic arterioles.

Other information about the role of adenine substances in controlling blood flow has come from gross and microscopic studies of the heart (Katori and Berne (40); Martini and Honig (41)) and from biochemical studies of smooth muscle (Honig (42)). From these studies a mechanism has been proposed for the manner in which the vascular smooth muscle, especially the precapillary sphincters,reacts to control blood flow through capillaries. The theory lends support to the hypothesis that the endothelial cells in the precapillary arteriole act as oxygen 'sensors' for the underlying smooth muscle cells, which are in intimate membrane contacts with endothelial cells via their myoendothelial junctions (Rhodin (21)).[1] The control concept of Honig

[1] For an alternative possibility, see page 92.

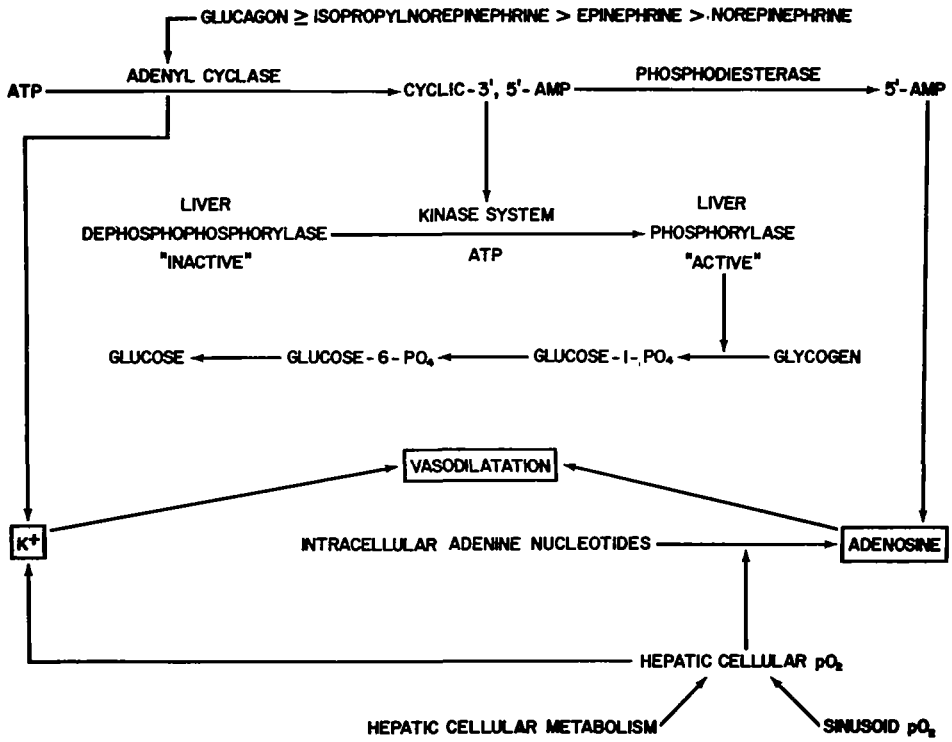


Fig. 3. Schema for Regulation of Blood Flow in the Liver (McCuskey (22))

(42) is briefly as follows: When the precapillary sphincter is open, oxygen transport does not limit mitochondrial function (in smooth muscle). The intracellular inorganic phosphate (P_i) and AMP attain their maximum concentrations, thereby releasing the inhibition of the contraction coupled ATPase of actomyosin. This permits sufficient tension to develop in smooth muscle to overcome the distending forces of the blood pressure. As the sphincter muscle shortens, the wall tension and load decrease due to the reduction in the vessel radius. Therefore, the muscles gain a progressively increasing mechanical advantage and the sphincter closes. Closure changes the diffusion distances and renders the muscles dependent upon bulk tissue oxygen. When the tissue metabolism lowers the pO_2 below 15-20 mm Hg, the mitochondria are unable to match the rate at which ATP is utilized. The resulting inhibition of ATP utilization decreases sphincter strength and the vessel opens. With blood flow reestablishment, mitochondrial respiration proceeds at maximum rates with a resulting decrease in P_i and AMP concentration which removes the inhibition of actomysin and restores contractile strength and the cycle is repeated. Aspects of the biochemical events of this theory have been found by in vitro analysis of non-vascular smooth muscle and support has been given from the observing the reaction of blood vessels in living cardiac muscle. The biochemical events need not be limited to smooth muscle because of the possibility that actomysin probably occurs in endothelial cells (Becker and Murphy (43)). Thus evidence is being secured

by various methods which is beginning to indicate some possible biochemical mechanisms that participate in the contractile and control processes of small blood vessels. The above mechanisms, when they are established, would help to explain some of the responses that are observed for controlling blood flow. Many problems remain. One of these is the source(s) for the intermittency of flow which is exhibited in mammalian skeletal muscle and frog liver. To what extent or to what degree intermittency of flow exists in other organs and species is still a moot question. For example, Richards and Schmidt (37) demonstrated a very striking intermittency of flow in glomeruli of frog kidneys. They observed flow and no flow states in glomeruli which were influenced by the extent of hydration and blood volume. While they unequivocally demonstrated that such marked intermittency could occur, the question that they did not answer was whether or not such striking changes occurred during normal physiological activities of the organ and animal. Their failure to do so produced doubts, as other investigators could not repeat their findings (Tamura et al.(44); Ellinger and Hirt (45); Grafflin and Bagley (46)). These investigators found that intermittency of flow decreased and 'disappeared' as the preparations of their specimens improved. Thus from such data it must be deduced that 'marked' intermittency of blood flow through the glomerular tuft of the frog kidney can and does occur but it only occurs when marked shifts in blood volume or other aberrations of physiology occurs. (As yet it has not been possible to observe the dynamics of blood flow in mammalian glomeruli in situ, but studies of blood flow in mammalian glomeruli of allografts in hamster kidney have demonstrated little if any intermittency of the type described by Richards and Schmidt (37) in these anesthetized animals (Oestermeyer, unpublished). These data and deductions must not be interpreted to mean that intermittency of glomerular blood flow does not exist in the mammalian kidney under physiological activity but rather that better experiments need to be designed to establish the dynamics of blood flow in glomeruli.)

Intermittency of blood flow through the hepatic lobule in the frog on the other hand has been repeatedly observed (Knisely, Bloch and Warner (29); Bloch (32); McCuskey (22); Debaker (31)). Such changes in flow have also been observed in mammalian hepatic lobules (Irwin and McDonald (30); McCuskey (22); Bloch (32)). Nevertheless the reaction is not pronounced in anesthetized mice (Bloch, unpublished). It is evident that more work must be done to establish the extent and stimuli that are responsible for producing intermittency of flow in the functional unit of the mammalian liver. Furthermore, almost nothing is known about such responses of the functional units in the brain, spinal cord, and gastrointestinal tract.

From the studies of functional units that have been made using the method of in vivo microscopy, it has become evident it is not sufficient to identify those components of the microvascular system such as arterioles, precapillary sphincters, capillaries, or venules that modify blood flow without securing simultaneous information of the functional state of the organ in which the unit is being investigated, as well as major functional cardiovascular and biochemical reactions of the entire animal. Unless this is done, the control mechanisms that are responding in the functional unit may be misinterpreted in respect to the complex responses exhibited by the organ at the time that the observations are made. (Methods are available that permit the securing of the requisite data.) An example can be cited.

It is almost universally assumed by investigators that the key role in controlling blood flow through the functional unit of mammalian skeletal muscle resides in precapillary sphincters. Teleologically it is a logical assumption. This concept was tested by examining a hypothesis that was proposed by Iberall (47) which stated that flow in nutrient channels among capillaries should be expected to be oscillatory and probably independent of the heart rate. The expected source for the control of the oscillations was considered to be the precapillary sphincters, or opening and closing of capillaries. It was possible to test the hypothesis in mammalian muscle of unanesthetized mice by utilizing a transparent chamber which permitted adequate resolution of terminal arterioles, their derivative capillaries in addition to the striations of the skeletal muscle fibers (Cardon, Oestermeyer and Bloch (15)). In brief, it was found that oscillations in red cell flow did exist that were independent of the heart rate but neither the terminal arterioles, precapillary sphincters, capillaries nor immediate post-capillary venules participated in the reaction. Instead the periodic fluctuations consisted of the number of red cells passing through a capillary segment. Not any of these morphological components, especially the precapillary sphincters, participated in regulating blood flow through the functional unit of skeletal muscle even when oxygen concentrations of the inspired gases varied from 20 to 8 per cent for periods ranging from 2 to 8 hours. However, transient disturbances of red cell oscillations in flow occurred which were directly proportional to the hypoxia. (The source for the oscillations are still to be determined.) The experiment does illustrate that modification will be required of the concept which states that the endothelial cell is a 'sensor' of oxygen which is responsible for initiating contraction of the smooth muscles of the precapillary sphincter.

Intermittency of flow in the vessels of the microvascular system are an expression of changes in the magnitude and distribution of blood volume. Therefore, we may ask what is the role of the microvascular system in controlling blood volume. The question will be considered in respect to the normal control of blood volume that are induced by physiological functions which form the spectrum of normal activity (e.g., exercise, postprandial digestion, seasonal variations in external temperature) rather than pathological adjustments of blood volume (e.g., hemorrhagic or traumatic shock).

According to gross physiological experiments, moment-to-moment adjustments of blood volume are accomplished by variations in the diameter of venous units (the capacitance vessels). These alterations in diameter are supposed to occur in venules whose diameters are greater than approximately 200 micra and in veins (e.g., Gauer and Henry (3)). While venules with a diameter of 200 micra are part of the microvascular system, such vessels, in mammals, are not a component of the functional units of organs, and therefore, in gross physiological data, are often overlooked as participating in controlling the blood volume, from moment-to-moment. Thus, at the present, the extent, if any, of the participation of the microvascular system of functional units in controlling blood volume from moment-to-moment, that is of immediate and moderate shifts in blood volume, is open to questions, as is the participation of venules whose diameters measure less than 200 micra.

The problem that requires an answer is the extent of the participation of blood reservoirs in controlling blood volume under normal physiological functions as described above. (The term, blood reservoir, is often used to describe the capacitance vessels plus the vessels of the liver, spleen and bone marrow. The term is used here to denote vessels that can withhold blood from the circulation for various periods of time; such vessels exist only in the liver, spleen and bone marrow, namely the sinusoids.) The reservoir function of the spleen has been studied extensively in experimental animals, especially in the dog and cat, and in man the organ has been studied primarily in disease. The blood storage of the human liver can be great in disease, but there is some question to what extent this organ and the spleen participate in controlling volume flow under normal physiological functions where 'moderate' or 'slight' alterations or shifts in blood volume are required. (See Gauer and Henry (3).) The liver in frogs, mice, rats, guinea pigs, rabbits and monkeys possess the structural components (as determined in the living liver) which can control flow through each sinusoid and thereby each sinusoid becomes a 'minute transfusion bottle'. The sinusoids of the spleen can react in a similar manner. (Too little is known as yet of the dynamics of the functional unit of living bone marrow but preliminary studies indicate that these sinusoids are also capable of being blood reservoirs (Kinosita (45), McCuskey (49)). Thus it is likely that the functional unit for storage (e.g., the venous system, others) requires careful definition.

The functional units of the liver are considered to participate in controlling blood volume, after the initial adjustments of the volume have been made by the venous system, the capacitance vessels. However, as the functional units of the liver control the blood flow to the right heart from the gastrointestinal tract and spleen, these units probably respond almost as soon as the capacitance vessels of the lower extremity. (In short, the liver is an extension of the right heart. Furthermore, the functional units of the liver also control the cell to plasma ratio, since the sinusoids can add packed cells to the systemic circulation.)

The rate at which volume redistribution under normal physiological requirements occurs between organs is essentially unknown. The rate is anticipated to be irregular as the adjustment would be one of organ demand (physiological activity). The rate of the shift in volume would be reflected in the intensity of the function, that is, be proportional to the volume demand. The suggestion that has been made by Iberall (47) that such a shift[1] might be accounted for by a seven minute resistive cycle might be possible, but it is not reasonable because of the location of the principal blood reservoir, the liver, and the temporal functions of metabolic demands.

The sensors for initiating the control response are probably from many sources and are not solely the baroreceptors in the thorax (e.g., Gauer and Henry (3)) or those associated with large veins; possibly the sensory nerves that accompany all blood vessels may act as such sensors. It is in-

[1] Iberall's suggestion referred to a shift in the volumes associated with the extremities.

triguing to speculate that at least some of the sensory nerves are pressure or volume detectors for the venous system either as physical or chemical transducers.

In summary, the role of the functional units of organs in regulating blood volume in regard to slight or moderate increases or decreases in volume is essentially unknown, but it is anticipated that the capacitance vessels and hepatic units are probably involved.

Finally,while many problems are unanswered in regard to the participation of the functional units in controlling blood volume, there is an even larger area of unknown factors in regard to the control of vascular permeability. As indicated, numerous vasoactive substances affect the walls of small blood vessels including capillaries, sinusoids and immediate post-capillary venules, but apparently not any one substance is the primary 'permeability factor' as considered by Krogh and Lewis. The magnitude of the problem for transvascular transit may be appreciated by the fact that there is little unanimity of opinion in regard to the cellular sites in the wall of capillaries or sinusoids where the transit of large molecules (i.e., approx. 1,000 mol. wt. or greater) occurs. Almost every component of the walls of these vessels has been considered at one time or another to participate in controlling the transit of large molecules, but to date there is a considerable gap between morphologist and physiologist for transvascular transit, let alone the factors that control such transit. Certainly everyone would subscribe to the paramount need for oxygen and glucose and the removal of wastes, but this does not provide too much basis for control factors or mechanisms.

The operational characteristics of the microvascular system require that they be analyzed in specific functional units in the living system using direct microscopy when at all possible as the primary detector. While numerous characteristics of the functional units can be examined with few corollary methods for providing information about the entire organ in which a functional unit is examined by microscopy, information about the status of the whole animal becomes necessary when the control of functional units are to be determined. The minimum additional information that should be secured should pertain to gross cardiovascular and nervous systems. Considerable progress can be made in the immediate future by using existing methods that will further elucidate the functional characteristics of major organs (see below). It will require but little additional effort, but considerably more support in material and personnel, to secure biophysical and chemical data simultaneously from functional units, the organ that the units are embedded in, and from the entire organ and animal.

A method that makes the functional unit available for study by direct microscopy in the living animal, are the various modifications of the transparent chamber methods. (In many instances the results obtained by these methods can be checked by direct microscopy of an organ in situ by using the quartz rod (Knisely) or by the Bloch method of transillumination).

The rabbit ear chamber method permits the examination of the simplest microvascular system, namely that of connective tissue, in an unanesthetized

animal. Modifications of this type of chamber permit the examination of some major organs, sometimes in an unanesthetized animal. For example, a chamber can be inserted into the abodminal wall of an animal (rabbit, dog, monkey) wherein a loop of intestine can be examined with the microscope in an unanesthetized animal, at least in the rabbit and dog. The chamber has also been modified so that the functional units of skeletal muscle can be examined with the microscope in an unanesthetized animal (mouse or rat). This was accomplished by placing the chamber in a dorsal skin fold of the rat or mouse. The subdermal layers in these animals contain striated muscle so that the responses of this major organ can be examined (Algire (50)). Finally, recently there has been developed a chamber method for the cheek pouch of the hamster. The importance of this chamber does not lie in regard to the cheek pouch but rather that allografts of major organs thrive in chambers of this animal. The specific major organ that thrives, rather uniquely, in this chamber is the kidney (Sanders (51); Greenblatt et al. (52); Oestermeyer (53)). With the establishment of a growing allograft of kidney, it has become possible, for the first time, to examine the microvascular system of glomeruli with adequate microscopic resolution, as well as other components of the nephron. Thus, the chamber methods provide an opportunity to examine functional units of some major organs, often in an unanesthetized animal, with adequate resolution, under sterile conditions for days, weeks or months (Williams (54)) and can often be compared with the responses of the functional units of the organs in situ (Bloch (7)).

A great advantage is secured when good optical images are obtained, as such images can then be transduced by such electronic image processing devices as television systems (Bloch (55)). When such image transduction is used in conjunction with monochromatic light, the images can be analyzed by the methods of scanning absorption or emission spectroscopy. When such a combination of methods are applied to the functional units not only is temporally sequential data of their morphokinesis secured but information of biochemical events that can also be secured by such methods (e.g.,oxidation reduction reactions). Such image analysis can be made of entire functional units or parts of such units as small as approx. 0.5 micra. The rate of the image production of the television system is 1/30 second, and when any of the approximately 490 horizontal lines that compose that image is selected, it can be analyzed at a rate of 53 msecs. (Bloch (55)). While the collection of such data forms the basis for analyzing a functional unit, additional information is required when their control, in relation to other organs, is to be assessed. Also, general information is required from the cardiovascular system and the nervous system, so as to assess the general physiological status of the animal, and in addition, specific biophysical and biochemical data should be secured that pertain to the particular aspect of the control system that is being studied. Such studies may and often do require the insertion of catheters and other probes to secure samples from the cardiovascular, urinary or nervous systems. When such manipulative procedures are required, it is usually not technically feasible to do this, plus microscopy, in a small animal (mouse) simply because of physical limitations. Furthermore, too frequently, the methods of gross physiology and biochemistry are not applicable to small animals when multiple samples must be secured at frequent intervals. Therefore, it is desirable to use a chamber method in a

large animal and an animal whose physiology and biochemistry has been thoroughly examined, to wit, the dog. In such animals a chamber can be inserted which contains skeletal muscle that would permit adequate microscopy and the affluent and effluent bloods to the muscle could be monitored and analyzed by using indwelling catheters, thereby securing the requisite samples required for determining gross and microscopic responses simultaneously. The type of data that could be secured from the microscopic images of the functional unit of skeletal muscle would be the dynamic morphology of the unit, and when spectroscopy is used, the oxidation-reduction reactions of both myoglobin and hemoglobin. Such data would be securable at repetition rates of 1/30 second and could be continued for hours. Such data could then be correlated with respiratory functions, as influenced by varying the oxygen concentration of the inspired air, or by stimulating the muscle under observation. The effluent blood could be analyzed for oxygen, carbon dioxide, lactic acid, glucose, or specific hormones that are considered to be important in muscle metabolism.

A wide variety of stimuli should be applied to a functional unit to determine to what extent its microvascular system is 'controlled'. The base line of the control for the microvascular system, in experimental animals, occurs when the animal is unanesthetized and quiescent but unrestrained, a state of physiology where minimum activity and presumably control exists. A variety of stimuli can be used to modify the responses of the local and whole cardiovascular system (e.g., by varying temperature, stimulation of specific sites in the central nervous system, etc.) in the healthy animal and these responses should be compared with the responses that occur in disease (see below).

As the nervous system has pronounced influence on the cardiovascular system and presumably on the microvascular system of the functional units of major organs, experiments should be conducted to assess the effects of the nervous system. Fortunately the chamber methods provide conditions wherein the direct effects of the nervous system can be examined. Allografts do not have direct connections to the host's nervous system for a period of probably several weeks. Alternately, a tissue like skeletal muscle can be studied with intact nervous connections and then it could be denervated by physical or chemical means.

The effects of oxygenation of the functional units of organs via their microvascular system can be assessed by controlling the available respiratory oxygen, by altering the viscosity of the blood by producing anemia, polycythemia or cellular (erythrocyte) aggregation. The latter mechanism can be produced by using high molecular weight dextrans or by disease. Malaria is a particularly useful disease as it can alter the blood in two ways; by the parasitemia which destroy the hemoglobin (e.g., in Plasmodium knowlesi malaria of monkeys, 7-900 erythrocytes or more per 1000 rbc's may be parasitized (Knisely et al.(56)), or the viscosity of the blood can be altered by red cell aggregation. The aggregation or the parasetemia can be controlled with chemotherapeutic substances; thereby the two effects of the disease on the functions of the blood can also be examined separately.

In conclusion, it is possible to examine the microvascular system in functional units of some major organs with methods that permit the securing of their dynamic morphology, and when such methods are used in conjunction with electronic image processing devices and monochromatic light, then the methods of spectroscopy permit the securing of biochemical information. The chamber methods are particularly applicable for elucidating some of the control mechanisms of the microvascular system of functional units of major organs. When chambers are inserted into 'large' animals then it will become possible for the microphysiologist to join forces with other disciplines like gross physiology and pharmacology so that the complete spectrum of at least certain dynamic responses can be evaluated simultaneously in the same animal at the same time. It is anticipated that such an approach should be fruitful for all concerned.

Addendum 1

Microvascular Geometry

As a useful bridge between the microvascular and gross circulatory systems (to be discussed later), we can summarize the geometric and topological characteristics of the arterial system in mammals. This system not only spans the spatial spectrum from the macrovasculature to the microvasculature, it also spans the temporal spectrum from genetic, developmental, and adaptive time scales to the operational time scale that reflects the operationally responsive characteristics of the circulation (i.e., its beat by beat character). The adaptive time scales extend from years down to the period of weeks. The operational time scale extends up to the period of a day. (e.g., a mammal does not complete its cycle of motor performance in less than a day, for minimally it must include the rest-wake cycle.) In this description, we shall not be concerned with the dynamic characteristics of the system, only with an 'average' description that is independent of time, independent of mammalian species, and independent of organ. By this very choice, the description is necessarily crude, but it attempts to be general.

The model is assembled from the following sources: Jeffords and Knisely (11), Patel et al. (57), Suwa (12), Mall (27), Iberall (58). It can be broken down into the following elements.

(a) Taper in the arterial system. This may be seen quite clearly in the work of Jeffords and Knisely (11), Patel (57), and in Suwa (12). Patel shows an average 23 kg. living dog's aorta. It tapers with a vertex angle of about 1.8^{o}. The data of Knisely's for tubes with diameters in the range 20-500 micra (both rat and frog) show tapers with vertex angles in the range 0.1^{o} to 2.3^{o}, averaging 1.0 ± 0.3^{o}. We will show later that these tapers are consistent with branching laws.

At the capillary level, the taper in a capillary segment (i.e., between bifurcations) is essentially zero.

On the venous side, the tapering begins to go the other way.

(b) Branching in an arterial level. Suwa's data suggest that it is not likely that a tube will branch in less than 3 diameters. When the branching does occur, the first few branches will only lead off a modest fraction of the tube area. There will probably be further side branching for a stretch of about 20-30 diameters. It appears likely that a tube will produce an equal bifurcation by 30 diameters. At the aorta, Patel shows a bifurcation at 20 diameters.

Thus on the average, one surmises that an arterial level will likely extend about 25 ± 5 diameters before it disappears as an equal bifurcation to a subsequent level. There will be approximately 8 branches per level (i.e., nominally 25 diameters divided by 3 diameters). However, these branches will not be equal in diameter.

We may illustrate the scatter from Patel's aorta data, and suggest it as 'typical' for a single arterial level.

Patel's Aortic Branches (57)

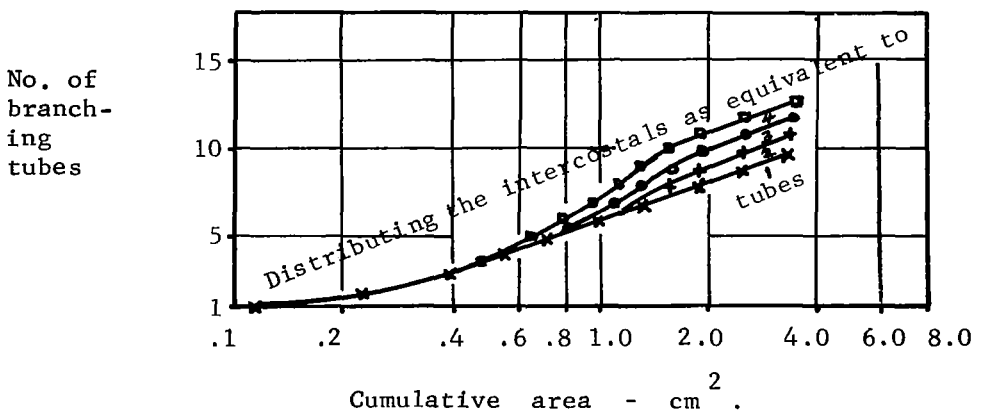
(These are major arteries in increasing order of size.)

Branch Areas cm^2
.876
.609
.353
.273
.173
.173
.170
.115
.115
20 x .029 = .580
3.437 cm^2

Area of aorta entrance 3.2 cm^2

We note the large number (20) of intercostals. Thus the number of branches per level might be taken as 29 (counting individual intercostals), or 10 (if the intercostals were counted as one equivalent tube).

If we regard the collection of these tubes to be normally distributed (as tested in their cumulative character), we find that the intercostals should be treated as equivalent to 1, 2, or 3 tubes, i.e., that the aorta has 10 to 12 branches. Since the numbers are comparable, as an approximate average, we will take 8 major branches per level for all levels, and as a further crude approximation, we will consider that they are all the same size. For greater accuracy, we might assume that the branchings for any level scatter 'normally' over a four to one range in diameter.



(c) Branching sizes in an arterial level. It was shown by both Groat ((59) and personal communication) and Suwa, that as a result of branching at a variety of levels

$$d_0^{2.7} = d_1^{2.7} + d_2^{2.7} + \ldots$$

where d_0 is the inlet diameter, and d_1 are the branching diameters. However, this law is limited to the range of diameters 20-25 micra to 1000 micra.

We may judge from Patel's data on the aorta, from Mall's, Thom's, and Bencke's data (27), that efflux area is conserved for larger tubes, say above 1 mm. in diameter, i.e.,

$$d_0^2 = d_1^2 + d_2^2 + \ldots$$

We can apply the branching rules to note the taper angle, if uniform. If a constant (equal to the average) branching diameter

$$d_0^n = N\, d_1^n$$

$$\frac{d_1}{d_0} = \frac{1}{N^{\frac{1}{n}}}$$

n = the exponent, such as 2, or Groat-Suwa's 2.7.
N = the number of assumed equal branches per level.

However, the last bifurcation leads to

$$2\, d_1^n = d_2^n$$

$$\frac{d_2}{d_1} = 2^{\frac{1}{n}}$$

$$\frac{d_2}{d_0} = \left(\frac{2}{N}\right)^{\frac{1}{n}}$$

Let θ = vertex angle

$$\theta = \frac{d_0 - d_2}{L} = \left(\frac{d}{L}\right)_o \frac{d_0 - d_2}{d_0}$$

$$= \frac{d}{L}\left[1 - \left(\frac{2}{N}\right)^{\frac{1}{n}}\right]$$

Let L/d = 25, N = 8

$$\theta = \left[\frac{1 - \frac{1}{4^{1/2}}}{25} \right] = 1^{o} \text{ for } n = 2,$$

$$= \left[\frac{1 - \frac{1}{4^{1/2.7}}}{25} \right] = 0.9^{o} \text{ for } n = 2.7$$

One may note that the taper is of the correct order of magnitude, the estimate being better for the small tubes below 1 mm.

Tables summarizing the characteristics of arterial trees have been prepared by Green (60), and Iberall (58). Green's table was based on Mall's (1888) data derived from a mesenteric artery in a small dog. Iberall's table was based on Mall's (1888-1906) data derived from a mesenteric, a stomach, a hepatic, a splenic, and an adrenal artery in the same small dog; Patel's data averaged from a considerable number of dog's aortas; Suwa's data derived from arteries in the basal ganglia, kidney, pancreas, heart, hind limb, intestine, cerebral cortex in the human; Groat's studies in a number of circulations; and from Adachi's atlas. It is the extensive number of organ circulations that have been quantitatively explored (in dog and human) that provides some confidence in 'universal' branching laws.

Let us apply the branching laws first to a 23 kg. dog, which starts from an aorta entrance diameter of 2.0 cm.

L/d = 25, N = 8

Level	Av. entrance diam.-cm.	No. tubes
1	2.0	1
2	0.71	8
3	0.25	64
4	0.088	510
5	0.041	4100
6	0.019	33,000
7	0.0088	260,000
8	0.0041	2.1×10^{6}
9	0.0019	17×10^{6}
10	0.0009	13×10^{7}
11	0.0004	1.1×10^{9}

The last line, the capillary line, is only an order of magnitude estimate for the number of capillary 'shrubs' or 'tufts' that branch off from a metarteriole. Each such 'shrub' or 'tuft' is a complex series-parallel array that differs from organ to organ, and thus leads to a much larger number of capillary segments (that is of 'runs' between bifurcations).

Addendum 2

Transcapillary Exchange

Basic questions that must be answered at the functional system level are the nature of metabolic and 'communicational' exchanges that occur between the units of the functional system. One of the primary exchange paths is transport across the capillary walls. While some of the basic 'facts' with regard to gross organ exchange have been established (e.g., Pappenheimer), and the nature of the exchange has been modelled at the level of the capillaries (Starling, Krogh, Landis, Renkin), the actual 'proof' at the microcirculation level has been harder to come by. We will illustrate the problem with regard to one particular aspect - filtration of water.

In rudimentary summary, it is believed that material - molecules and larger entities - is transported across membranes in general, and past or through endothelial cells in capillaries in particular. With regard to molecules, it is believed that small molecules pass with much higher filtration coefficients than larger molecules. (The logarithm of the filtration coefficient is believed to drop linearly but precipitously with increasing log of molecular size.) Above about 40 Å molecular size, it is believed that the filtration coefficient remains more nearly constant and low. This has been explained by a 'pore' theory, in which a small percentage of the endothelial wall (say 0.1% of the area, in particular at the junction between cells) has 'pores' or cracks of about 40 Å unit radius or opening. Calculations and estimates have typically been made on the basis of gross bulk material flow laws, such as Poiseuille flow.

The transport of larger sized molecules is currently explained as a vesicular transport in which the large molecules are engulfed in the membrane wall and pass across the endothelial cell (Bruns and Palade, II, (61)). In fact, Bloch (62) has shown that large sized particles up to 1-2 micra also pass through the endothelial cell.

An 'explanation' of net water filtration through capillaries was provided by Starling, by the following mechanism. Essentially all ingredients in the watery environment are the same on both sides of the capillary, except for the presence of a higher concentration of large protein molecules in the plasma. This would create an inward osmotic drive of water from the tissue to the capillaries as it tried to dilute the plasma. However, there is an equivalence between hydrostatic drive due to the pressure head, and the osmotic drive.[1] Thus at the entrance of capillaries, with higher hydrostatic head, the water filtration was outward from the capillaries, but at the exit of capillaries with higher osmotic head, the water would be returned and 'reabsorbed'.

[1] The equivalence is confirmed in modern irreversible statistical thermodynamics by virtue of the Onsager relations.

The ubiquity, the essential requirement for, and the complexity of water filtration makes it an apt microvascular problem to consider.

The current physiological textbook descriptions of filtration may be examined in the following references:

Ruch and Fulton (1960)
Ponder (revised) - pp. 754-756, 758-763.
Hernández-Peón - pp. 878-880.

Ruch and Patton (1966)
Catchpole - pp. 621, 624-628.

Best and Taylor (1961)
Gregg - pp. 157, 158.

Guyton (1966)
pp. 434-443.

Davson (1964)
pp. 287-289.

Landis-Pappenheimer-Renkin[1] analysis (63)

Poiseuille flow in a pore:

$$\Delta p = \frac{128}{\pi} \frac{q\mu l}{d^4}$$

q = flow per pore
l = length of flow channel

Suppose N pores per unit capillary wall area

Flow area of the holes = $\frac{\pi}{4} d^2 N = fA$

A = capillary area per unit tissue volume (or mass).
f = fraction of the total capillary area occupied by pores.
Q = total filtration flow per unit tissue volume (= NAq)

$$\Delta p = 32 \frac{Q}{NA} \frac{\mu l}{\left(\frac{\pi}{4} d^2\right) d^2} = \frac{32 Q \mu l}{A^2 f d^2}$$

K_f = filtration coefficient $\left(= \frac{Q}{A\Delta p}\right)$

$$K_f = \frac{d^2}{32\mu} \frac{fA}{l}$$

[1] Renkin's current data taken from summary lecture at the 1970 Microcirculation Society meeting.

He assumes the following constants:

$A = 7000\ cm^2/100$ gm. tissue (the specific capillary wall area).
$l = 1.3$ micra
$f = 0.001$
$fA/l = 55{,}000$ cm/100 gm. tissue.
$d = 80$ Å (40 Å radius pore).
$\mu = .01$ poise (gm/cm sec.) - water viscosity

$$K_f = 1.1 \times 10^{-7} \frac{cm^5}{dyne\ sec.} \ / \ 100\ gm.\ tissue.$$

According to the Starling hypothesis:

$$Q = K_f A \left[p_i - p_o - \Pi_i + \Pi_o \right]$$

p_i = hydrostatic pressure inside of capillaries.
p_o = hydrostatic pressure outside of capillaries.
Π_i = colloid osmotic pressure inside of capillaries.
Π_o = colloid osmotic pressure outside of capillaries.

$\Pi_i = 25$ mm. Hg. (human serum).
$\Pi_o = 5$ mm. Hg. (uncertain).
(Wiederhelm has more recently proposed 9 mm. Hg.)
$\Pi_i - \Pi_o = 20$ mm. Hg.
$p_o = 2$ mm. Hg.
p_i (arterial end) = 32 mm. Hg.
p_i (venous end) = 15 mm. Hg.
$p_i - p_o$ (arterial) = 30 mm. Hg.
$p_i - p_o$ (venous) = 13 mm. Hg.
net pressure (arterial) = 10 mm. Hg. (outward filtration)
(venous) = -7 mm. Hg. (inward reabsorption)

With some uncertain scatter on the components, this is the classic Starling filtration-reabsorption model for water filtration, and the Pappenheimer-Renkin pore model.

We wish to reexamine it in the light of two repeats by Zweifach (64) of Landis' experiment of occluding a capillary and noting the direction and magnitude of plasma efflux, as indicated by the coasting of the red cells toward or away from the occlusion.

First, with regard to parameters for the model.

(a) Capillary area. We estimate - for muscle - an average capillary separation among parallel muscle fibers of 60 micra, or about 200 capillaries per mm^2. However, we estimate in longitudinal passage, a length of about 100 micra. Thus there are 2000 capillary 'segments'[1] per mm^3. At each run, we estimate that there are about another 200 capillary connections per layer. With ten layers, this is about 2000 additional capillary segments for a total

[1] A 'segment' of capillary is that run between bifurcations.

of 4000 capillary segments per mm^3. (We estimate that this is for normal resting muscle, and that for muscle at maximum exercise, there are twice as many, or 8000 capillary segments per mm^3. On the other hand, Martin et al. (65) and others suggest numbers that are 5-10 times larger than these.)

Thus the capillary area, assuming 6 micra average capillary, is 7500 cm^2/ 100 gm. tissue at rest, or 15,000 cm^2/100 gm. tissue at peak activity. Our estimate thus agrees with Renkin's.

(b) Path length. Thicknesses here are conventional. One may choose numbers, for endothelial thickness, from 0.2 to 2 micra. Thus it is the effective fA/l which is pertinent. Pappenheimer et al.(1951) indicates 120,000 cm. per 100 gm. tissue; Renkin (1970) prefers 55,000 cm. per 100 gm. One notes that this corresponds to f = 0.1% of the area.

At this point, we should note a first criticism. Bruns and Palade (I, (61)), after conducting an electronmicroscopic study of capillaries in muscle, conclude:

"Isolated, perfused legs of small mammals ... have been used in studies of capillary permeability. To explain permeability data obtained on such preparations, Pappenheimer et al ... postulated ... pores (water-filled channels) which have an effective radius of ~30Å, an aggregated area of 0.1 to 0.2% ... Following an earlier suggestion of Chambers and Zweifach ... the pores were assumed to be the intercellular spaces ... More recently, Landis and Pappenheimer ..revised the effective pore radius upward to 45Å, and assumed that the pores are actually ... the intercellular spaces because the width (100A) and the ... (intercellular space) area ... estimated from available electron micrographs ... were found to correspond to those expected for pores.

"In the material examined, we did not find pores of the dimensions, distribution, and location postulated. The fractional surface ... of intercellular spaces ... calculates at 0.1% ... all intercellular spaces ... should be open to a gap of ~100Å.

"Pores penetrating ... the ... periphery of the cell should be easily visible ... None has been found."

Thus 'pores', apparently are fictitious 'effective' functional structures, rather than fixed geometric structures (i.e., walls act as if they possessed pores).

(c) Filtration. Zweifach (1966) found that serum seemed to move out of the capillaries observed in 85% of the cases.

We thus turn to his second round of studies (64). In rabbits (rabbit omentum)

$$\Pi_i = 19 \text{ mm. Hg.}$$
$$K_f = 0.008 \mu \text{/mm Hg. sec.}$$

In all of the capillaries observed, whether on the arterial side or venous side, he found an outward filtration of plasma. "We have not, to date, either in the omentum of the rabbit, or in the mesentery of the rat, encountered vessels which consistently show inward filtration."

We find, as a summary of his data

$$\Pi_o - p_o = 5 \text{ mm Hg.}$$

$$\frac{Q}{A} = K_f \left[p_i - \Pi_i + \Pi_o - p_o\right]$$
$$= .008 \left[24 \text{ mm.} - 19 \text{ mm.} + 5 \text{ mm.}\right] = .08 \text{ micra/sec.}$$

$$\text{Average observed } \frac{Q}{A} = 0.06 \text{ micra/sec.}$$

For arterial capillaries

$$Q/A = .008 \left[30 - 19 + 5\right] = 0.13 \text{ micra/sec.}$$

For venous capillaries

$$Q/A = .008 \left[18 - 19 + 5\right] = 0.032 \text{ micra/sec.}$$

Zweifach notes that no experimental points have been obtained in the region in which the osmotic pressure exceeds the hydrostatic pressure.

(We would tend to use a slightly different set of values

$$Q/A = .008 \left[20 - 19 - 1 + 7\right] = 0.06 \text{ micra/sec. average}$$
$$Q/A = .008 \left[24 - 19 - 1 + 7\right] = 0.10 \text{ micra/sec. arterial}$$
$$Q/A = .008 \left[16 - 19 - 1 + 7\right] = 0.024 \text{ micra/sec. venous}$$

The difference is due to the fact that we have chosen different values for entrance and exit hydrostatic pressures in capillaries, as a summary of data in the literature. Nevertheless the filtration is outward.)

This filtration coefficient 0.008 micra/sec. mm Hg is equal to 45×10^{-7} $\frac{cm^5}{\text{dyne sec.}}$ per 100 gm.

This is 30 times Renkin's value.

Zweifach (1970) also studied the cremaster muscle in the rat,

$$K_f = 0.0014 \text{ micra/sec. mm Hg}$$
$$p_i = 24 \text{ mm Hg. (arterial); } = 16 \text{ mm Hg (venous).}$$

Let's examine his balance

$$.006 \text{ micra/sec. (observed)} = .0012 \frac{\mu}{\text{sec mm Hg.}} \text{ (observed) } \left[24 \text{ mm.} - 25 - 1 + 7\right]$$

on arterial side.

Zweifach points out that if reabsorption took place, it had to take place beyond the primary venule, i.e., beyond 12 micra diameter venules.

The filtration coefficient of 0.0014 micra/sec. mm. Hg. is about five times as large as Renkin's value.

From this study, we draw the following conclusions: A 'filtration' coefficient, for a given molecular species, is not a constant, but may easily vary over one or even two decades. A near 'constancy' of an effective pore size cannot be a geometric parameter, but if anything, a dynamic parameter. Water moves out in the capillary bed. If it returns by the Starling mechanism, it has to return in the venous bed. Zweifach's data suggest that the filtration constant may in fact increase on the venous side.

If we consider transport of oxygen, a different material, there is recent evidence (Duling, Berne (66)) that perhaps oxygen is appreciably transported through the arteriolar bed before reaching capillaries. While the conclusion is inadequately proven, nevertheless it is one further instance that suggests that the specific dynamic nature and locus of material and chemical transit in the microvasculature is still an open subject for study.

Bibliography

1. Green, H. D., and Kepchar, J. H., Control of Peripheral Resistance in Major Systemic Vascular Beds, Physiol. Rev., 39, 617, 1959.

2. Folkow, B., Nervous Control of Blood Vessels, Physiol. Rev., 35, 629, 1955.

3. Gauer, O. H., and Henry, J. P., Circulatory Basis of Fluid Volume Control, Physiol. Rev., 43, 423, 1963.

4. Guyton, A. C. and Coleman, T. G., Long Term Regulation of the Circulation: Interrelationships with Body Fluids, p. 179, In "Physical Bases of Circulatory Transport: Regulation and Exchange," Ed. by Reeve, E. B. and Guyton, A. C., W. B. Saunders, Co., Philadelphia, 1967.

5. Haddy, F. J. and Scott, J. B., Metabolically Linked Vasoactive Chemicals in Local Regulation of Blood Flow, Physiol. Rev., 48, 688, 1968.

6. Mellander, S. and Johansson, B., Control of Resistance Exchange and Capacitance Functions in the Peripheral Circulation, Pharmacol. Rev., 20, 117, 1968.

7. Bloch, E. H., The Hepatic Microvascular System, In "The Living Microvascular System," Ed. by Bloch, E. H., W. B. Saunders Co., Philadelphia, 1970 (in press).

8. Heymans, C. J. F. and Folkow, B., Vasomotor Control and Regulation of Blood Pressure, p. 407, In "Circulation of the Blood: Men and Ideas," Ed. by Fishman, A. P. and Richards, D. W., Oxford Univ. Press, New York, 1964.

9. Zweifach, B. W. and Metz, D. B., Selective Distribution of Blood Through the Terminal Vascular Bed of Mesenteric Structures and Skeletal Muscle, Angiology, 6, 282, 1955.

10. Clara, M., "Die Arterio-Venösen Anastomosen," Springer, Wien, 1956.

11. Jeffords, J.V. and Knisely, M. H., Concerning the Geometric Shapes of Arteries and Arterioles, Angiology, 7, 105, 1956.

12. Suwa, N., Niwa, T., Fukasuwa, H. and Sasaki, Y., Estimation of Intravascular Blood Pressure Gradient by Mathematical Analysis of Arterial Cells, Tohoku J. Exper. Med. 79, 168, 1963.

13. Nicoll, P. A. and Webb, R. L., Vascular Patterns and Active Vasomotion as Determiners of Flow Through Minute Vessels, Angiology, 6, 291, 1955.

14. Honig, C., Frierson, J. and Patterson, J., Comparison of Neural Controls of Resistance and Capillary Density in Resting Muscle, Am. J. Physiol, 218, (4), 937, 1970.

15. Cardon, S., Oestermeyer, C. and Bloch, E., Effect of Oxygen on Cyclic Red Blood Cell Flow in Unanesthetized Mammalian Striated Muscle as Determined by Microscopy, Microvasc. Res., 2, 67, 1970.

16. Funaki, S., Studies on Membrane Potentials of Vascular Smooth Muscle with Intracellular Microelectrodes, Proc. Jap. Acad., 34, 534, 1958.

17. Steedman, W. M., Micro-electrode Studies on Mammalian Smooth Muscle, J. Physiol. (Lond.), 186, 382, 1966.

18. Siggins, G. R., Evidence for Reciprocal Innervation of Arterioles in the Frog Retrolingual Membrane, Bibl. Anat., 9, 328, 1967.

19. Johansson, B. and Ljung, B., Spread of Excitation in the Smooth Muscle of the Rat Portal Vein, Acta. Physiol. Scand., 70, 312, 1967.

20. Su, C., Bevan, J. A. and Ursillo, R. C., Electrical Quiescence of Pulmonary Artery Smooth Muscle During Sympathomimetic Stimulation, Circ. Res., 15, 20, 1964.

21. Rhodin, J.A.G., The Ultrastructure of Mammalian Arterioles and Precapillary Sphincters, J. Ultrastruct. Res., 18, 181, 1967.

22. McCuskey, R. S., A Dynamic and Static Study of Hepatic Arterioles and Hepatic Sphincters, Am. J. Anat., 455, 1966.

23. Ahlquist, R. P., Taylor, J. P., Rawson, C. W. and Sydow, V. L. Comparative Effects of Epinephrine and Levarterenol in Intact Anesthetized Dog, J. Pharmacol. and Exper. Therap., 110, 352, 1954.

24. Ahlquist, R. P., Adrenergic Drugs, In "Pharmacol. in Medicine," Ed. by Drill, V. A., McGraw-Hill, New York, 1958.

25. Green, H. D., Comparison in Man of Adrenergic Blockade Produced by Dibenzyline, Ilida, Priscoline and Regitine, Circulation, 15, 47, 1957.

25a. Mitchell, G. A. G.,"Cardiovascular Innervation," E. S. Livingston, Ltd., Edinburgh, 1956.

26. Jones, H. W., Hoerr, N. L. and Osol, A., "Blakiston's New Gould Medical Dictionary," The Blakiston Co., Philadelphia, 1949.

27. Mall, F., Königs. Sächs, Ges. Wiss. Abhandl.,Math. Cl., 14, 151, 1888. Amer. J. Anat., 5, 227, 1906.

28. Bloch, E. H., The Termination of Hepatic Arterioles and the Functional Unit of the Liver As Determined by Microscopy of the Living Organ, N. Y. Acad. Sci., 1969 (in press).

29. Knisely, M. H., Bloch, E. H. and Warner, L., Selective Phagocytosis, Kgl. Danske Videnskab. Selskab. Biol. Skrifter, 4, 1, 1948.

30. Irwin, J. and MacDonald, J., Microscopic Observations of the Intrahepatic Circulation of Living Guinea Pigs, Anat. Rec., 117, 1, 1953.

31. Debaker, H. S., "Microscopic Observations of the Hepatic Circulation in Living Frogs and Mice," PhD. Thesis, Medical College of the State of South Carolina, Charleston, South Carolina, 1967.

32. Bloch, E. H., The In Vivo Microscopic Vascular Anatomy and Physiology of the Liver as Determined with the Quartz Rod Method of Transillumination, Angiology, 6, 340, 1955.

33. Barcroft, H. and Swan, H. J. C., "Sympathetic Control of Human Blood Vessels," Edward Arnold & Co., London, 1953.

34. Krogh, A., "The Anatomy and Physiology of Capillaries," Yale Univ. Press, New Haven, 1922, 1929.

35. Shepherd, J. T., "Physiology of the Circulation in Human Limbs in Health and Disease," W. B. Saunders Co., Philadelphia, 1963.

36. Lewis, T., "The Blood Vessels of the Human Skin and their Responses," Shaw and Sons, London, 1927.

37. Richards, A. N. and Schmidt, C. F., A Description of the Glomerular Circulation in the Frog Kidney and Observations Concerning the Action of Adrenalin and Various Other Substances Upon It, Am. J. Physiol., 71, 178, 1924.

37a. Ebbecke, U., Die Lokale Vasomotorische Reaction (LVR) der Haut und die Innere Organe, Pflügers Arch., 169, 1, 1917.

38. Fulton, G. P. and Lutz, B. R., The Neuro-Motor Mechanism of the Small Blood Vessels of the Frog, Science, 92, 223, 1940.

39. Bloch, E. H., Some Actions of Adrenalin Chloride and Acetyl-Beta-Methyl Choline Chloride on Finer Vessels of Living Frog Liver Lobules, Anat. Rec., 76 7 , 1940.

40. Katori, M. and Berne, R. M., Release of Adenosine from Anoxic Hearts, Circulation Res., 19, 420, 1966.

41. Martini, J. and Honig, C. R., Direct Measurement of Intercapillary Distance in Beating Rat Heart in situ Under Various Conditions of O_2 Supply, Microvasc. Res., 1, 244, 1969

42. Honig, C. R., Control of Smooth Muscle Actomyosin by Phosphate and 5' AMP: Possible Role in Metabolic Autoregulation, Microvasc. Res., 1, 133, 1969.

43. Becker, C. G. and Murphy, G. E., Demonstration of Contractile Protein in Endothelium and Cells of the Heart Valves, Endocardium, Intima, Arterio-

sclerotic Plaques, and Aschoff Bodies of Rheumatic Heart Disease, Am. J. Path., 55, 1, 1969.

44. Tamura, K., Miyamura, K., Nishina, T. and Nagasawa, H., Studies in the Excretion of Urine, Jap. J. Med. Sci. (Trans.), (Pharmacol.), 1, 211, 1927.

45. Ellinger, P. and Hirt, A., Mikroskopische Untersuchungen an Lebenden Organen, Arch. Exp. Path. u. Pharmakol., 145, 193, 1929.

46. Grafflin, A. L. and Bagley, E. H., Glomerular Activity in the Frog's Kidney, Bull. Johns Hopks. Hosp., 91, 306, 1952.

47. Iberall, A. S., Study of the General Dynamics of the Physical-Chemical Systems in Mammals, NASA CR-129, 1964.

48. Kinosita, R., Studies on Bone Marrow Biodynamics, Bibl. Anat., 1, 106, 1961.

49. McCuskey, R. S., The Microvascular System of Erythropoitic Organs, In "The Living Microvascular System," Ed. by Bloch, E. H., W. B. Saunders, Philadelphia, 1970 (in press).

50. Algire, G. H., An Adaptation of the Transparent Chamber Technique to the Mouse, J. Nat. Cancer Inst., 4, 1, 1943.

51. Sanders, A. G., Microcirculation in Grafts of Normal and Malignant Tissue, J. Anat., 631, 1963.

52. Greenblatt, M., Choudari, K. V. R. and Sanders, A. G., Mammalian Microcirculation in the Living Animal. Microcirculation of Renal Homograft Transplants in Hamsters, Am. J. Path., 56, 317, 1969.

53. Oestermeyer, C. F., A Modification of the Algire Chamber Method, Anat. Rec., 157, 294, 1967.

54. Williams, R. G., Microscopic Studies in Living Mammals with Transparent Chamber Methods, Int. Rev. Cytol., 3,359, 1954.

55. Bloch, E. H., A Quantitative Study of the Hemodynamics in the Living Microvascular System, Am. J. Anat., 110, 125, 1962.

56. Knisely, M. H., Stratman-Thomas, W. K., Eliot, T. S. and Bloch, E. H., Knowlesi Malaria in Monkeys, J. Nat. Malaria Soc., 4, 285, 1945.

57. Patel, D., DeFreitas, F., Greenfield, J. and Fry, D., Relationship of Radius to Pressure Along the Aorta in Living Dogs, J. Appl. Physiol., 18, 1111, 1963.

58. Iberall, A. S., Anatomy and Steady Flow Characteristics of the Arterial System with an Introduction to Its Pulsatile Characteristics, Math. Biosciences, Vol. 1, (3), 375, 1967.

59. Groat, R., Fed. Proc., 7, (2), 45, 1948 (Abstract); personal communication.

60. Green, H., Circulatory System: Physical Principles, in "Medical Physics," Vol. 2, Yearbook Publ., 1950.

61. Bruns, R. and Palade, G., Studies on Blood Capillaries, I. General Organization of Blood Capillaries in Muscle, J. Cell Biol., 37, 244; II. Transport of Ferritin Molecules Across the Wall of Muscle Capillaries, 37, 277, 1968.

62. Bloch, E. H. and Coyas, S., The Transit of Large Molecules and Particulate Matter Across Individual Endothelial Cells ..., Angiology, 15, (8), 353, 1964.

63. Landis, E. and Pappenheimer, J., Chapter 29, In "Handbook of Physiology, Sec. 2: Circulation, Vol. 2," Ed. by Dow, P. and Hamilton, W., Amer. Physiol. Soc., Washington, D. C., 1963.

64. Zweifach, B., Microvas. Res., 1, 83, 1968; 2, 96, 1970.

65. Martin, E., Wooley, F. and Miller, M., Am. J. Physiol., 100, 407, 1932.

66. Duling, B. and Berne, R., Fed. Proc. (Abstracts), 29, (2), March-April 1970, Abst. 427.

VII. THOUGHTS CONCERNING AUTONOMIC NERVOUS SYSTEM CONTROLLING INFLUENCES

Since the functional system at the microscopic level consists of the actuated or actuating element in the organ, its exchanging local blood supply, and its nerve elements, we must show concern with the lowest functional level of the autonomic nervous system. The lateral extent is still microscopic, yet it reaches out to span all of the organ systems.

In fact it is likely that the functional unit is best to be defined as the total organ element served by a single, sympathetic and parasympathetic plexus.

Classical anatomic and histochemical studies have established our general knowledge of the autonomic innervation of the various peripheral organs. The recent histofluorometric studies for catecholamine containing structures has opened up areas of study that were not possible with the conventional silver staining techniques. Since 1962 a great deal of new and unusual information has accumulated in this highly specialized literature which is not yet generally appreciated. The nature of the localization of sympathetic nerves in both the CNS and the periphery has revealed the sites of action of the sympathetic nerve neurotransmitter. Several unusual and unorthodox considerations of physiologic actions are presented below. The speculative nature and the departure from established doctrine in the discussion to follow are acknowledged by the writers. The purpose of this exercise, however, is, like most highly speculative discourses, to stimulate thought and experimental testing of ideas. Focus of attention aimed at those new developments that suggest peripheral autonomic controlling influences.

1. The Autonomic Ganglia

The ganglia are essentially intermediate relay stations that monitor electrical pulses emanating from the spinal cord and are thus mediators between impulses from the CNS and peripheral organs. The sympathetic ganglia, located along the sympathetic chain (para vertebral-ganglia) or within more peripherally located sites (pre vertebral-ganglia), spread information from the control nervous system diffusely to a number of organs. For example, one preganglionic fiber terminating in the superior cervical ganglion (SCG) makes synaptic contact with approximately 20 sympathetic ganglion cells which in turn send impulses to various organs such as the eye, salivary glands, thyroid, pineal, blood vessels of the head, neck and base of brain. A diffuse reaction can therefore be elicited. The ganglion can serve to regulate by repression or inhibition. Recent histofluorescence studies have demonstrated a pathway for ganglionic control (Norberg and Sjöqvist, 1966). An intraganglionic system of adrenergic terminals was described. These were subsequently shown to be axon collaterals from postganglionic processes that "feedback" upon the ganglion cells and possibly on presynaptic elements (Jacobowitz and Woodward, 1968). One can now believe that the sympathetic neurotransmitter may influence transmission through the ganglion (See Fig. 1).

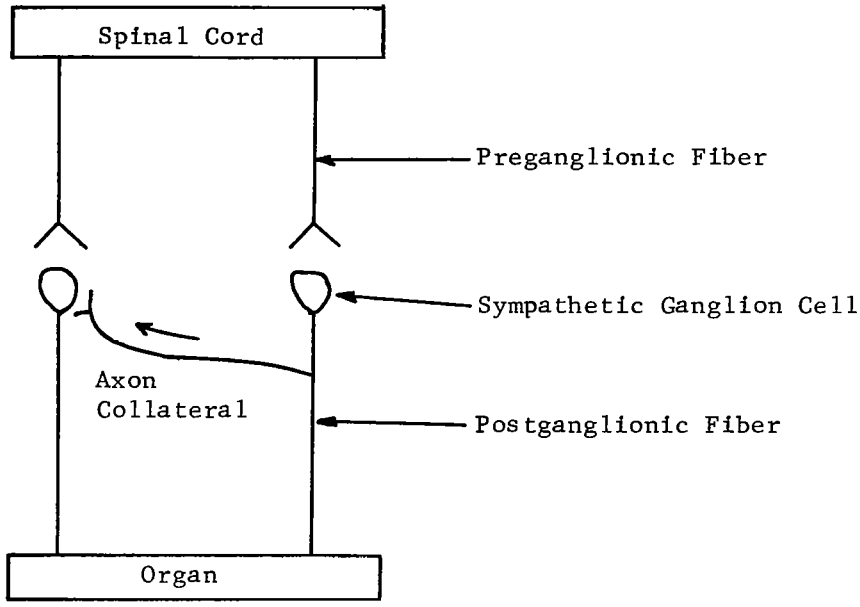


Figure 1

Recent pharmacologic studies (DeGroat and Volle, 1966) indicate that exogenous norepinephrine or epinephrine is capable of causing both depressant and facilitative actions on the cat SCG. These actions are mediated by the activation of alpha (depression) or beta (facilitation) adrenoceptive sites. The overall evidence would suggest that the sympathetic transmitter serves to influence, or modulate, transmission through the ganglion. Whether facilitation or inhibition results may very well depend upon such factors as the amount of transmitter released or the receptor site (pre- or postsynaptic) most sensitive to norepinephrine, depending upon prevailing conditions. At this level, then, there exists the potential for dynamic regulation, by mediation, mainly by inhibition or release from inhibition, of the required manifold of rhythmic processes which make up the many biochemical chains in the organism.

Study of the ganglion offers insight into the probable functional character of the brain which in a sense is a collection of multiple interrelated and intercommunicating ganglia. The ganglia can thereby be viewed as isolated little brains displaced from the CNS to the periphery (or conversely as long range mediators of secretory processes). This idea is what led Winslow to suggest the term sympathetic ganglia, for there seemed to be enough "little brains" in the chain to account for the sympathy of parts, as the integration of visual responses was then called. Our knowledge of the wiring within the ganglia is a first step to our approach to the understanding of the influences

of the adrenergic neurons within the brain. The recent disclosure of monoaminergic paths within the CNS, greatly extended through the work of Fuxe and others, has opened up new and fruitful fields of the highest significance.

2. Autonomic Action and Interactions at the End Organ

(a) Heart. - The microanatomic description of the ganglion, however, does not give us an insight into the living dynamic moment-to-moment language of the nervous system which results in long term oscillatory properties upon which we are focusing attention. Adequate information regarding electrical impulse traffic to and from ganglia in the unanesthetized animal is lacking, for technical reasons. However, if such information, probably possessing a high frequency oscillatory character, were available, there is some doubt whether this would reflect the autonomic activity at the synaptic level within the end organs. The release of norepinephrine at sympathetic endings may not bear any simple relation to the number of nerve action potentials traveling down the postganglionic fibres, but indeed the organ itself may signal to the numerous varicose nerve terminals and modulate the release of transmitter. If this speculative idea were true, it would constitute a microcontrol or a fine adjustment without CNS intervention. The manner in which such local adjustment could take place at smooth and cardiac muscle and glands could involve an alteration of the extracellular milieu in contact with the varicosity of the adrenergic terminals, which are the terminal storehouses for the neurotransmitter norepinephrine. It is proposed that a change in pH, pO_2 or pCO_2 created by the innervated end organ, may influence the terminal nerve membrane in such a manner that quanta of norepinephrine can be released which, in turn, may act upon the adrenoreceptor for the prescribed action. This concept evolved from a study of the adrenergic innervation of calves and steers in heart failure (Vogel, Jacobowitz, and Chidsey, 1969). A marked change in fluorescence intensity (measuring norepinephrine content) in the adrenergic innervation of the failing heart, was noted predominantly in the terminal varicose fibers in close approximation to muscle bundles and not in fibers coursing through connective tissue or around blood vessels. This indicates that fibers innervating the cardiac muscle were influenced in some manner by an alteration taking place at the region of close synaptic association between the nerve terminal and the cardiac muscle cell. This dramatic effect resulting from a stress situation (failing heart) is interpreted as a magnification of a subtle controlling influence of a coupling from muscle to nerve which may be a physiologic mechanism whereby moment-to-moment feedback control can take place without CNS involvement. One possible biochemical explanation for the mechanism of this muscle-nerve modulation could involve a metabolite, perhaps lactic acid, that is released in increased quantities in working muscles. An increased acidic environment at the nerve membrane can cause alteration of the active transport involved in the reuptake mechanism for norepinephrine so that greater release of the neurotransmitter ensues.

(b) Gut. - The feedback control by adrenergic axon collaterals of sympathetic ganglia cited above appears to be a general phenomenon that occurs in both sympathetic (preganglia) and parasympathetic ganglia that are located within the peripheral organs. (Localization of the latter ganglia within the

organ partly explains why parasympathetic effects, in contrast to those of the sympathetic, are often localized, i.e., are confined to a single organ or to a part of an organ.) Recent studies of the innervation of the gut demonstrated that the primary adrenergic innervation of the alimentary tract is at the myenteric plexus of Auerbach, i.e., the sympathetic innervation is not to the smooth muscle of the gut as classically believed, but to the parasympathetic ganglia situated between the muscle layers of the entire alimentary canal (Jacobowitz, 1965). The cholinergic innervation is comparatively dense. This general picture of the autonomic innervation suggests that the contraction of the gut is due to the release of acetylcholine from the endings of postganglionic fibers terminating on muscle fibers. The adrenergic inhibitory mechanism therefore takes place primarily at the myenteric ganglia (Fig. 2). A pharmacologic correlate to support this hypothesis has been repeatedly observed in studies in which administration of norepinephrine or epinephrine to ganglia produces a block in synaptic transmission. Here then is the first demonstration of sympathetic control at the organ level of an intermediate relay station, the parasympathetic ganglion.

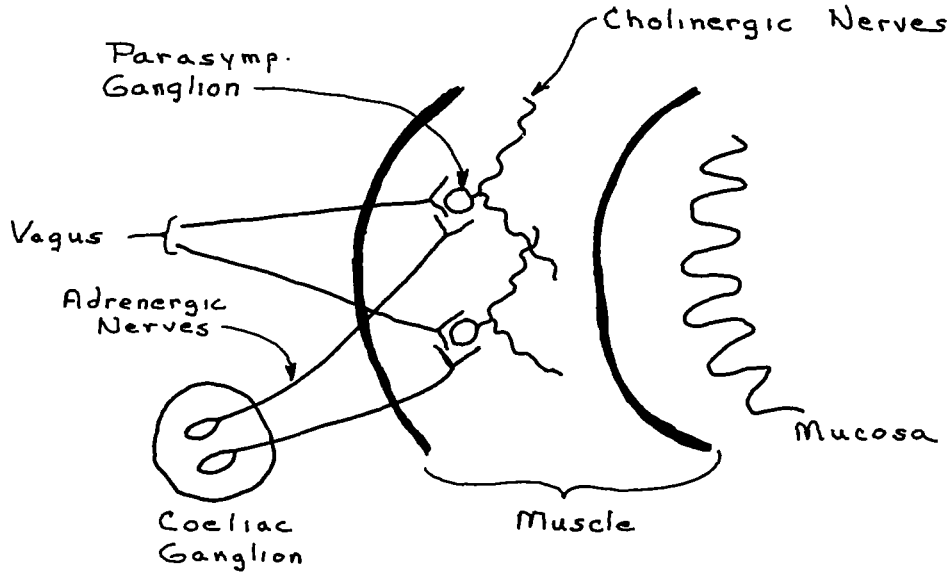


Figure 2

It is of more than neuroanatomic interest that sympathetic nerve fibers (in addition to parasympathetic and sensory fibers) enter organs in proximity to and around blood vessels. In the gut it has been observed that adrenergic fibers emanate from the surface of arteries (tunica adventitia) within the muscle

layer and the submucosa and terminate at the Auerbach's plexus ganglion cells. It is suggested that peristaltic activity of the gut, in attempting to advance a bolus of food or feces, creates a pressure upon the arteries (containing the sympathetic nerve plexus whose cell bodies are located outside the gut); thereby causing a local nerve discharge, or action potential, which serves to inhibit ganglion impulse traffic through cholinergic nerves; thus resulting in relaxation of the gut at the bolus level (See Fig. 3).

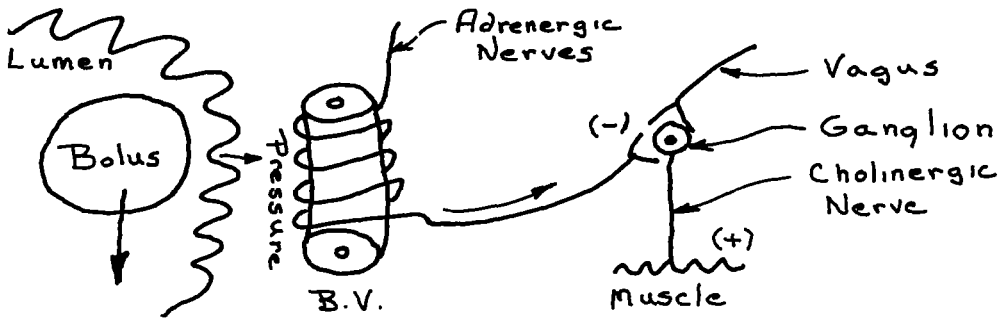


Figure 3

Here, the blood vessel is visualized as a strain gauge. We are aware that pressure applied to nerves can cause an electric discharge. The pressure, then, is caused by the bolus of material transmitting force through the mucosa and submucosa to the plexus of nerves surrounding the vasculature, creating local nerve action potentials, resulting in relaxation of the gut through inhibition of parasympathetic nerve transmission through the ganglia. This is somewhat analogous to the axon reflex described in sensory nerve receptors or sympathetic terminals (Coon and Rothman, 1940) in the skin. The axon reflex is a response in which nerve impulses initiated in sensory or autonomic nerve terminals (by injury or perhaps pressure) are relayed antidromically down other branches of the nerves. The axon reflex in autonomic nerves is an area of study that has been totally neglected. It is a phenomenon thay may be capable of local control of autonomic functions without CNS intervention.

The sympathetic nerve ending is no longer considered to be a single anatomic terminal in apposition with a muscle or gland cell, but rather numerous physiologic endings contained within an autonomic ground plexus. Discharge of norepinephrine molecules occurs simultaneously throughout the network of varicose terminals in response to a nerve action potential that traverses the preterminal process. There is no reason to believe that the impulse traffic is unidirectional in the adrenergic nerve plexus. A stimulus, such as pressure, may readily cause an impulse discharge along the localized plexus of nerve terminals.

(c) Pineal. - Another area where sympathetic axon reflex-like action may be taking place is the eye-pineal gland axis. After histochemical observations of the sympathetic ganglion cells of the superior cervical ganglia of several species it became clear that many of the neurons are multipolar with axons that bifurcate frequently (Jacobowitz and Woodward, 1968). In addition, previous studies have demonstrated that there was a significant increase in the norepinephrine content in the posterior segment of the eyes of albino guinea pigs and rabbits exposed to light for 1 hour (Nichols, Jacobowitz, and Hottenstein, 1967). This suggested that light may act directly on the nerve terminals to cause an increase in synthesis of choroidal and iris norepinephrine. This fact taken together with the multipolar nature of the neurons of the superior cervical ganglion suggested a mechanism whereby light information could be transferred from the eye to the pineal gland, i.e., by conduction of impulses from adrenergic nerves within the eye to the pineal by way of neuronal branches from the same neuron emanating from the same cell body within the superior cervical ganglion. This concept is diagrammatically shown in Fig. 4.

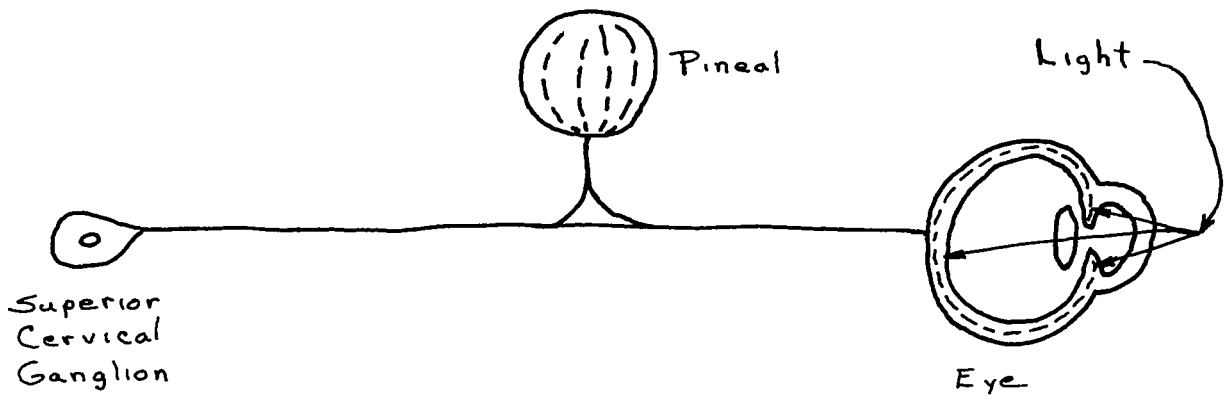


Figure 4

Previous studies have shown that the rat pineal gland exhibits a diurnal variation in its content of serotonin, norepinephrine and the enzyme hydroxyindole-O-methyl transferase((HIOMT) which is responsible for the conversion) for the synthesis of melatonin and that each of these rhythms is affected by changes in environmental lighting (Quay, 1963; Axelrod, Wurtman and Snyder, 1965; Wurtman et al., 1967). In addition, superior cervical ganglionectomy abolishes all of these rhythms. It was therefore proposed that each of these rhythms is maintained by neural information transmitted to the pineal by an unknown pathway from the CNS through the cervical sympathetic trunk and finally through the superior cervical ganglion. From a study designed to investigate whether individual neurons within the superior cervical ganglion were capable of sending axons to both the eye and pineal, a working hypothesis was developed such that light information could be transferred from the eye to the pineal (Jacobowitz, et al., 1968). It is suggested that information about

lighting is mediated by conduction of impulses from choroidal and iris adrenergic terminals to similar terminals in the pineal via fibers emanating from a system of multipolar neurons within the superior cervical ganglion. It can be further suggested that conduction of impulses from the eye to the pineal takes place by antidromic impulses. This in effect is analogous to a sensory system as indicated diagrammatically (Fig. 5).

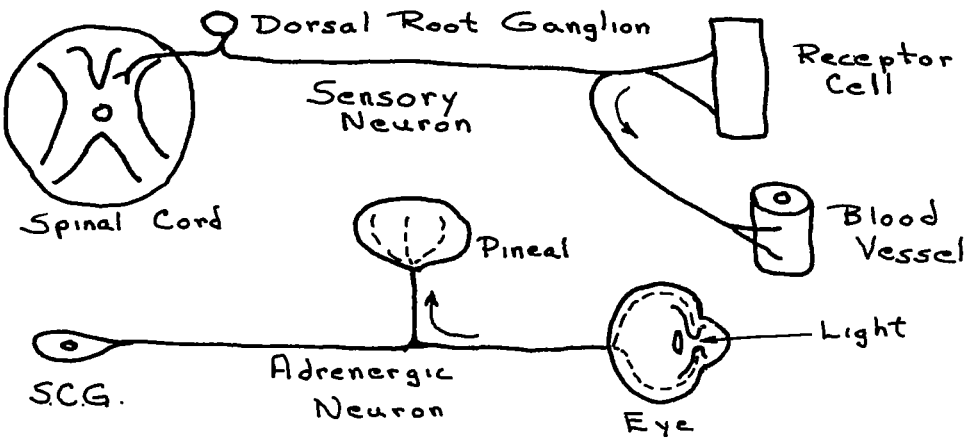


Figure 5

The superior cervical ganglion would be analagous to the dorsal root ganglion; the adrenergic nerve terminals would be analagous to the receptor terminals and the pineal would be analagous to the dorsal root of the spinal cord.

This system can also be viewed to be analagous to an axon reflex, a response in which impulses initiated in sensory and autonomic nerves by a stimulus are relayed antidromically down other branches of the nerve plexus causing a reflex response such as vasodilation.

The suggestion that autonomic impulses take an antidromic path is admittedly most unorthodox and may appear to challenge established principles underlying our knowledge of the nervous system. However, no studies have been made within tissues that establish the direction of impulse traffic in the autonomic ground plexus of the autonomic nervous system. On the other hand pharmacologic studies made with intraarterial injection of acetylcholine into an organ have resulted in both release of norepinephrine within the organ and antidromic impulses along the sympathetic nerves emanating from the organ (Balkely et al., 1963; Ferry, 1968), thereby indicating that sympathetic antidromic impulses are possible.

Characterization of catecholamine-containing nerves as sensory fibers is not without precedence. Dahl et al., (1963) have shown that sensory adrenergic neurons occur commonly within widely separated invertebrate groups. For example, the tentacular nerve net of the sea-anemones consists of adrenergic varicose fibers and is believed to have a sensory or possibly a combined sensory and motor function.

The above influence of light on the pineal norepinephrine may indeed affect the daily rhythms of various substances in this gland. "Almost all daily rhythyms in mammals appear to be generated by 'biologic clocks' originating within the animal. These mechanisms operate with a periodicity approximating, but not equalling, 24 hours; hence they have been termed 'circadian'. Such rhythms are ordinarily entrained by the external day-night cycle of environmental lighting; however, they persist when animals are blinded or are kept in continuous darkness. An exception is several measured cycles in the rat pineal.[1] This appears to depend entirely upon the animal's ability to perceive rhythmic changes in environmental lighting. All circadian rhythms studies up to this stage had in common the ability to persist for some weeks after animals were deprived of environmental lighting cues (by blinding or being placed in darkness). These rhythms no longer showed a period of precisely 24 hours, but they did fall in a range between 22 and 26 hours and hence were thought to be regulated by some internal mechanism not dependent on, but usually synchronized with, environmental lighting." (Wurtman and Axelrod, 1965). The mechanisms for circadian rhythms, however, are still unknown. Within the pineal, catecholamines are offered as a possible chemoelectric escapement mechanism.

3. The Cardiovascular System

The autonomic nervous system is prominent among the control mechanisms available to the body for the modulation of cardiovascular performance. The sympathetic nervous system is particularly important when the metabolic requirements of the body necessitate levels of cardiac work which approach maximal. In any discussion of the autonomic innervation of the heart we learn that the parasympathetic nerves carried to the heart by way of the vagus, carry impulses only to the sinoatrial and atrioventricular nodes and to the atrial myocardium (not the ventricle). In contrast, sympathetic innervation is said to be distributed to all areas of the circulation, including the myocardium and specialized conduction structures of both the atria and ventricles and the smooth muscle of the walls of the arterioles. Discussion of the innervation of the veins is scanty and confusing.

(a) Cholinergic nerves. - Recent unpublished studies of the cholinergic innervation of the heart showed that a marked species variation of the presence of cholinergic nerves to the ventricles exists. A considerable number of cholinergic nerves was noted in the ventricles of the dog, pig and ewe whereas in smaller animals, like the rat, mouse and rabbit, no nerves were present.

[1] These likely refer to variations in serotonin, 5 HIOMT, and norepinephrine.

Demonstration of cholinergic nerves in the ventricular myocardium would support and explain on a neuroanatomic basis the observation of physiologists who show that stimulation of the vagi may cause significant negative inotropic effects on the dog left ventricle (DeGeest et al., 1965; Levy and Zieske, 1969). To complete our knowledge of the cholinergic innervation to the heart it should be noted that the atrioventricular conducting system of the heart (Purkinje fibers) within the ventricle is heavily innervated with cholinergic nerves. This fact does not seem to be generally appreciated.

(b) Adrenergic nerves. - Recent studies of the sympathetic innervation of the heart show that in addition to the sympathetic innervation to the myocardium, adrenergic nerves were observed in contact with the cholinergic ganglion cells within the heart which are in the vicinity of the SA and AV nodes within the interatrial septum (Jacobowitz, 1967). Since sympathetic ganglion cells are not normally present in the heart, it is reasonable to assume that these adrenergic nerve terminals probably arise from sympathetic ganglia outside the heart. Such terminals were not seen in denervated cat hearts. As indicated above, considerable pharmacologic evidence indicates that catecholamines are capable of depressing ganglionic impulse transmission. By examination of intracellular electrical responses in single neurons of the parasympathetic ganglia of the frog heart, Topchieva (1966) showed that stimulation of the sympathetic nerve to the heart causes hyperpolarizing potentials indicative of inhibitory postsynaptic potentials. From this the author concluded "sympathetic nerve fibers as well as vagus fibers may end on neurons in the intramural ganglia of the frog heart".

(c) Chromaffin cells.

(1) Heart. - In addition to the terminal adrenergic network of nerve fibers within the ganglia, intense green-yellow fluorescent cells were seen both within and in the vicinity of the ganglia of all species studied (Jacobowitz, 1967). These cells have been identified as chromaffin cells. They are found most frequently within the cardiac ganglia in close proximity to ganglion cells, although they are also found outside the confines of the ganglion in the epicardium. They have striking resemblence to chromaffin cells of the adrenal medulla. They are often associated with rich vascular networks in a characteristic whorl-like fashion. They are also often seen within large nerve trunks. A network of smooth processes was occasionally seen to emanate from the catecholamine containing cells. These processes are extensions of the cells and come in close apposition to ganglion cells, blood vessels and adrenergic nerve trunks. Sympathetic denervation does not appear to alter the content or intensity of fluorescence of the catecholamine containing cells.

Further studies on the innervation of the chromaffin cells of the heart show that these cells receive a postganglionic parasympathetic innervation from the cholinergic ganglion cells. Since sympathetic preganglionic fibers are not found in the heart, the presence of cholinergic (acetylcholinesterase-staining) fibers, here, is a reflection of parasympathetic preganglionic (vagal) or postganglionic fibers. From a physiologic point of view, pre- or postganglionic innervation to the cardiac chromaffin cells would, in all probability, serve the same functional purpose. The possible physiologic significance

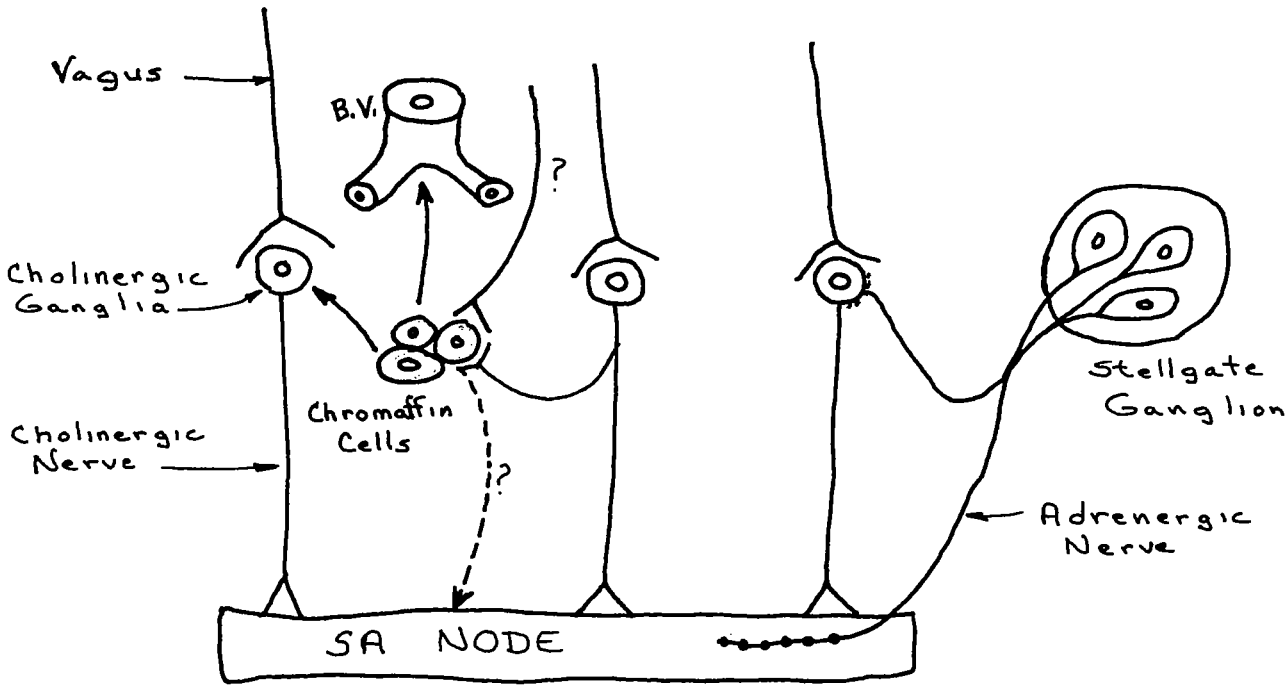


Figure 6

of the chromaffin cell is indicated diagrammatically in Fig. 6. The presence of chromaffin cell masses in or near the ganglion suggests an influence of the chromaffin cells on ganglionic transmission (designated by an arrow directed toward the ganglion cell). The processes that occasionally emanate from the chromaffin cells and are sometimes seen to be in close relationship to ganglion cells suggest a manner of catecholamine release from these cellular elements. The release of catecholamine upon ganglion cells should have the same effect as that suggested previously for the system of adrenergic fibers in close proximity to the cardiac ganglion cells, i.e., depression of ganglionic transmission. It is suggested that the chromaffin cells act to supplement the existing adrenergic innervation to subserve a modulatory influence on cholinergic transmission through the ganglion. The postganglionic cholinergic innervation to the chromaffin cell seems to be analogous to the Renshaw cell system postulated in the spinal cord, i.e., a motor neuron innervation resulting in a 'feedback' inhibition. Such a ganglion-chromaffin cell arrangement would suggest a self-controlling system within the heart. An unusual increase in vagal impulses could result in a release of catecholamine from both the chromaffin cells and adrenergic terminals (sympathetic reflex) to dampen the excessive cholinergic inhibition. Such an influence would provide support for the proposed sympathetic mechanism for vagal escape (Campos and Friedman, 1963). The localization of chromaffin cells around blood vessels suggests an influence on ganglionic vascular tone which could thereby affect impulse transmission.

In addition, release of catecholamine into ganglion capillaries could provide an additional route by which chromaffin-cell products are made available. Direct release of catecholamines on atrial muscle is questionable. Preganglionic vagal or sympathetic influence on chromaffin cells has not been established.

(2) Carotid and aortic body. - In addition to their location in the confines of ganglia, catecholamine cells are also found in the vicinity of the aorta. The greatest concentrations of these cells are found in the tunica adventitia between the aorta and pulmonary artery slightly above the level of the semilunar valves. These masses are often ovoid in shape and are arranged in loose cords surrounding small blood vessels and capillaries. These cells are more properly called aortic bodies. The comparable glomus cells of the carotid body has only recently been shown to contain large amounts of catecholamines. In addition, they give a positive chromaffin reaction and are essentially identical to the chromaffin cells of the heart and adrenal medulla. The carotid and aortic bodies are chemoreceptors which receive a rich sensory innervation. These bodies are stimulated by anoxia, hypercapnia and acidemia (like adrenal medulla chromaffin cells). The response of the chemoreceptors to anoxia is of vital importance to the body; after denervation of chemoreceptors, anoxia causes respiratory depression and death, whereas if they are intact, anoxia induces reflex hyperpnea and reflex cardiovascular excitation. The chemoreceptors are of almost negligible importance in the control of the circulation of an intact animal at rest. They become important, however, in conditions of circulatory insufficiency, anoxia, and although their activity during muscular exercise has never been directly investigated, it has been supposed that the chemical changes in the blood such as occur during heavy exercise would stimulate the chemoreceptors and hence the medullary cardiovascular centers (Heymans and Neil, 1958).

The chromaffin (or glomus) cells which constitute the aortic and carotid bodies have recently been shown to contain large amounts of catecholamine. (The human carotid body also contains serotonin within the glomus cells). These cells are the chemoreceptor cells. The function of the catecholamines, if any, in the initiation of the sensory discharge is unknown. The chemoreceptor (chromaffin) cells possibly release catecholamine that acts on the sensory nerve endings which lie in close apposition to the cells. Low concentrations of epinephrine infused into the carotid body have been shown to increase the sensitivity of the chemoreceptor to other stimuli such as acetylcholine. Large concentrations decrease the sensitivity to such stimuli. Much work needs to be done here. Currently, there are several investigators who believe that the transmitter substance is acetylcholine (Eyzaguirre et al., 1965) although this is refuted. The presence of large amounts of acetylcholine in sensory nerve endings is strange since we are presently unaware of the nature of the sensory nerve transmitter, if any. The carotid body does have a large number of acetylcholinesterase-staining nerve fibers (presumably cholinergic) enveloping the chromaffin chemosensory cells (Biscoe and Silver, 1966). Cutting the sinus nerve has no effect on the distribution of the cholinergic nerves within the cat carotid body, which indicates that acetylcholine is not localized in the sensory nerves. It appears that these cholinergic nerves are postganglionics from the superior cervical ganglion which contain a small

population of cholinergic ganglia (innervates sweat glands and blood vessels). The role of catecholamines and acetylcholine in the initiation of sensory discharge remains obscure. A possible unique relationship suggests itself. Just as cholinergic (preganglionic) nerves innervate adrenal medulla chromaffin cells, cholinergic (postganglionic) nerves may innervate the chemosensory chromaffin cells. In addition, the close proximity of the sensory endings to the chromaffin cells (shown by electron micrographic studies, Lever, et al., 1959) suggests that catecholamine release may alter sensory nerve impulse discharge. The following scheme is indicated (Fig. 7).

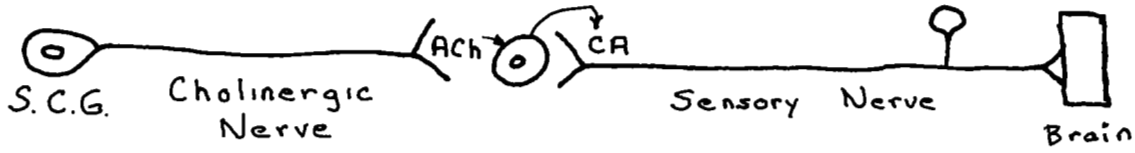


Figure 7

(3) Adrenal medulla. - This is analagous to the local 'feedback' system suggested by Nijima and Winter (1968) in the adrenal medulla of the rabbit and cat. They found afferent discharges emanating from nerves innervating the adrenal gland. They further proposed that there are chemosensitive receptors which constitute the afferent endings of a local feedback loop. Catecholamines and acetylcholine caused a depression of the spontaneous firing of the sensory nerve. We would suggest a system in the adrenal medulla analagous to that shown in Fig. 7 except that the cholinergic nerve would be preganglionic. Release of acetylcholine from the cholinergic endings would cause release of catecholamine from the chromaffin cells which would act upon the sensory nerve endings which would modify (depress) neuronal discharge rates as suggested by Nijima and Winter (1968). This, in effect, would serve as a local 'feedback' system which could moderate the flux of catecholamine secreted from the adrenal gland.

What begins to unfold in strategic areas of the body where control influences are important in maintaining the dynamic regulation of the living system are regions of catecholamine-containing chromaffin cells that may act as 'phase reversors' to achieve the desired 'inhibition or release from inhibition, of a manifold of oscillatory (or rhythmic) processes which make up the many biochemical chains in the organism'.

(4) Pancreas. - Another place where cells with large concentrations of either catecholamines or serotonin is found is in the islets cells of the pancreas. There is a great species variation here in terms of localization within either alpha (glucagon producing) or beta (insulin producing) cells or whether the amine is a catecholamine or 5-hydroxytryptamine (serotonin). Amine fluorescence was observed in alpha cells in the duck, in the young pigmented mouse, rabbit, cat and dog; in beta cells only in the adult guinea-pig and human fetus; and in alpha and beta cells in the pig. The functional significance of the fluorescent amines present in the islet cells is not known

(Cegrell, 1968). It was suggested that the amine is associated with the granules which are regarded as the storage site for insulin in the beta cells. It is tempting to suggest that the monoamine in the alpha cells is associated with the hormone glucagon which is generally accepted to be secreted by this cell.

There are also intrapancreatic cholinergic ganglia in the exocrine glandular portion of the pancreas (Alm et al., 1967). These ganglia contain chromaffin cells similar to that observed in the heart. Adrenergic nerve terminals were also contained around some of the cholinergic ganglion cells analogous to the arrangement in the heart and gut. Here too, control influences on ganglionic transmission is implied. Whether or not there is a cholinergic influence on islet cell secretion is unknown.

(5) Thyroid. - Similar cell systems have been found to occur in the parafollicular cells of the thyroid gland. The parafollicular cells (not the follicular cells) in the sheep thyroid contain serotonin (Falck et al., 1964). Although no catecholamines or serotonin can normally be observed in the mouse thyroid, there is strong evidence that amine mechanisms operate in the parafollicular cells of the mouse thyroid gland (Larson et al., 1966). We do not begin to know the possible function of these cells; influence on thyroid hormone release remains to be seen.

(6) Ganglia. - Finally, another area of most recent interest is the presence of masses of chromaffin cells in the sympathetic prevertebral ganglia (e.g., superior cervical, stellate, coeliac and mesenteric). It has been demonstrated by electron microscopy that there are preganglionic synaptic contacts upon the chromaffin cells. The chromaffin cells in turn synapse with dendrites of the adrenergic neurons. In addition, the chromaffin cells are found in close proximity to fenestrated capillaries or in clusters next to these vessels in a manner similar to that seen in the adrenal medulla. It has been suggested that the chromaffin cells in the sympathetic ganglia act as inhibitory 'interneurons'. In addition, a localized endocrine effect on the entire ganglion through release of catecholamines into the fenestrated capillaries might serve as part of a portal system (Siegrist et al., 1968; Matthews and Raisman, 1969). This anatomic arrangement of the chromaffin cells to the ganglia and the proposed inhibitory influence is essentially identical to those described previously in the parasympathetic ganglia of the heart (Jacobowitz, 1967). The anatomic arrangement can be diagrammatically visualized as follows in Fig. 8.

This would serve to supplement the intraganglionic 'feedback' adrenergic fibers described previously which could modulate impulse transmission through the ganglia (Jacobowitz and Woodward, 1968). Inhibitory effects on the ganglia can be demonstrated by stimulation of the preganglionic trunk at low frequency which then facilitates transmission of a second volley. The second response is greater because more fibers are recruited after the second shock. After a stronger stimulation, facilitation is reduced, which indicates that an inhibitory influence becomes involved (Dunant and Dolivo, 1967). This negative influence may be due to both the intraganglionic adrenergic 'feedback' fibers and chromaffin cell secretion of catecholamine.

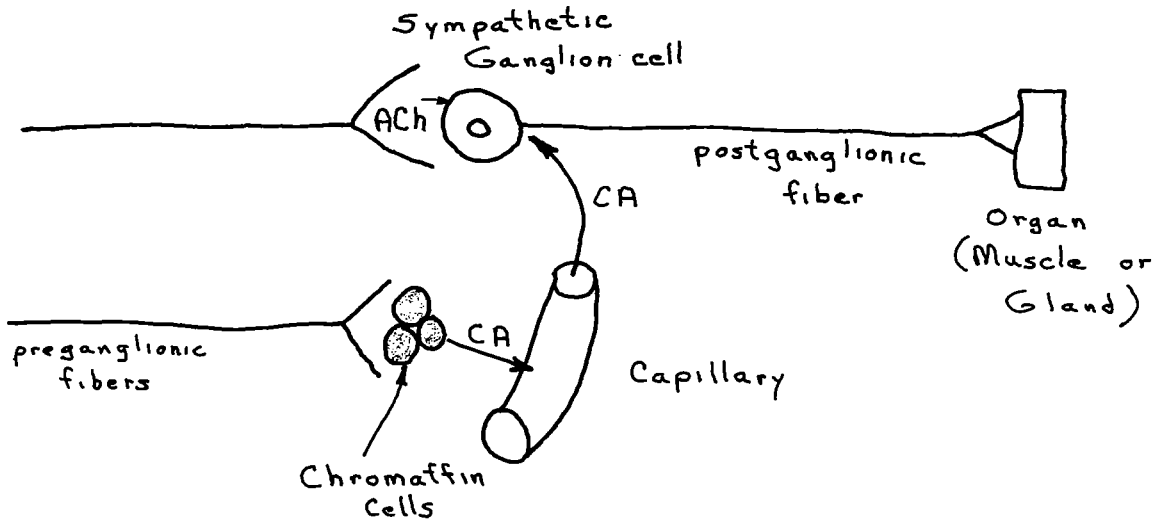


Figure 8

(d) Peripheral circulation. - An elegant review of control of peripheral circulation was recently published by Mellander and Johansson (1968). It is clear that there is still no real theory of circulation which would embrace all the relevant phenomena. There is no question that vessel-wall innervation plays an enormous part in the control of peripheral circulation. In the past, there have existed conflicting views on the problem of capillary innervation. It appears that most of the older literature describes nerves on capillaries. There was also no agreement on whether the nerves described were autonomic or sensory. The absence of agreement on capillary innervation was perhaps due to the fact that there was no generally agreed view on the structure of the capillary itself. Description of much of the capillary innervation was derived from observations of "arteriovenous bridges" containing muscle. Currently it is generally believed that the muscle-free true capillaries are devoid of a nerve supply. In recent times, innervation of vessels has been studied with the use of the catecholamine fluorescent method. It was shown that in most regions there is no evidence of any adrenergic nerve supply to the capillaries, even though adrenergic fibers occasionally appeared close to small unidentified vessels of capillary size (Ehinger et al., 1966). Adrenergic nerve fibers were, however, noted next to capillaries of the ciliary processes of the eye (Ehinger, 1966) and in the exocrine pancreas and islets of Langerhans (Alm et al., 1967). Using electron microscopic methods Ruskell (1969) describes innervation of both sympathetic and parasympathetic terminals to the capillaries within the lacrimal gland. It is of interest that, so far, innervation to capillaries has been described in glandular structures. Innervation to capillaries is strange since the classic teaching is that autonomic nerves can only innervate muscle (cardiac and smooth) and gland cells.

One suspects that the autonomic nervous system is not the simple and primitive affair of some imaginations, and its subtleties of arrangement and function are still largely unfathomed.

Bibliography

1. Alm, P., Cegrell, L., Ehinger, B., Falck, B., Z. Zellforsch., 83, 178, 1967.

2. Axelrod, J., Wurtman, R.J., Snyder, S.H., J. Biol. Chem., 240, 949, 1965.

3. Biscoe, T.J., Silver, A., J. Physiol., 183, 501, 1966.

4. Blakely, A.G.H., Brown, G.L., Ferry, C.B., J. Physiol., 167, 505, 1963.

5. Campos, H.A., Friedman, A.H., J. Physiol., 169, 249, 1963.

6. Cegrell, L., Acta Physiol. Scand. Suppl., 314, 1968.

7. Coon and Rothman, J. Pharmacol. Exptl. Therap., 68, 301, 1940.

8. Dahl, E., Falck, B., von Mecklenburg, C., Myhrberg, H., Quart. J. Micr. Sci., 104, 531, 1963.

9. DeGeest, H., Levy, M.N., Zieske, H., Lipman, R.I., Circ. Res., 17, 222, 1965.

10. DeGroat, W.C., Volle, R.L., J. Pharmacol. Exptl. Therap., 154, 1, 1966.

11. Dunant, Y., Dolivo, M., J. Physiol. (Paris), 59, 281, 1967.

12. Ehinger, B., Invest. Ophthal., 5, 42, 1966.

13. Ehinger, B., Falck, B., Sporrong, B., Bibl. Anat., 8, 35, 1966.

14. Eyzaguirre, C., Koyano, H.T., Taylor, J.R., J. Physiol., 178, 463, 1965.

15. Falck, B., Larson, B., Mecklenburg, C., Rosengren, E., Svenaens, K., Acta Physiol. Scand., 62, 491, 1964.

16. Ferry, C.B., J. Physiol., 167, 487, 1963.

17. Heymans, C., Neil, E., "Reflexogenic Areas of the Cardiovascular System," Churchill, London, 1958.

18. Jacobowitz, D., J. Pharmacol. Exptl. Therap., 149, 358, 1965.

19. Jacobowitz, D., J. Pharmacol. Exptl. Therap., 158, 227, 1967

20. Jacobowitz, D., Nichols, C.W., Saelens, J.K., Kovacsics, G.B., The Pharmacologist (abstract), 1968.

21. Jacobowitz, D., Woodward, J.K., J. Pharmacol. Exptl. Therap., 162, 213, 1968.

22. Larson, B., Owman, C., Sundler, F., Endocrinol., 78, 1109, 1966.

23. Lever, J.D., Lewis, P.R., Boyd, J.D., J. Anat., 93, 478, 1959.

24. Levy, M.N., Zieske, H., Circ. Res., 24, 303, 1969.

25. Matthews, R.J., Raisman, G., J. Anat., 105, 255, 1969.

26. Nichols, C.W., Jacobowitz, D., Hottenstein, M., Investig. Ophthal., 6, 642, 1967.

27. Nijima, A., Winter, D.L., Science, 159, 434, 1968.

28. Norberg, K.A., Sjoqvist, F., Pharmacol. Rev., 18, 743, 1966.

29. Quay, W.B., Z. Zellforsch. Mikr. Anat., 60, 479, 1963.

30. Ruskell, G., Z. Zellforsch. Mikr. Anat., 94, 261, 1969.

31. Siegrist, G., Dolivo, M., Dunant, Y., Foroglou-Keramous, C., De Ribaupierre, Fr., Rouiller, C., J. Ultrastruct. Res., 25, 381, 1968.

32. Topchieva, E.P., Fed. Proc., 25, Trans. Suppl. T739, 1966.

33. Wurtman, R.J., Axelrod, J.A., Scientific Amer., 213, 50, 1965.

34. Wurtman, R.J., Axelrod, J., Chu, E.W., Heller, A., Moore, R.Y., Endocrinol., 81, 509, 1967.

35. Vogel, H.K., Jacobowitz, D., Chidsey, C.A., Circ. Res., 34, 71, 1969.

VIII. THE CARDIOVASCULAR SYSTEM

1. Introduction

Blood flow, oxygen consumption, and metabolism are closely related. It is useful to characterize them as some of the fundamental properties of the mammalian CV system.

For example, because oxygen uptake and blood flow are correlated through the hemoglobin-carrier, and the hematocrit is nearly constant and because metabolic power production is very closely tied to oxygen consumption by virtue of the essential oxidative process, metabolism and blood flow should bear the same relation to animal size for all species (so that their ratio should be nearly independent of size).

We will attempt to characterize 'universal' systems relations for all (mature) mammals. By their very generality, they may be crude. Nevertheless, we will attempt to keep them centered, particularly within the range of dog to man. We want to assume that the relations are at least reasonable for the size range of shrews to large whales, or more restrictively at least the size range of rats to cattle.

(a) <u>A law of geometric development</u>. Topologically, regardless of the geometric shape, the body surface area seems to vary as the 2/3rds power of the body weight (or volume, assuming that all mammals approximately show the same density, that of water).

$$A = \left(\frac{A}{W^{2/3}}\right)_o W^{2/3}$$

$$A \left(\frac{\rho}{W}\right)^{2/3} = 9.5 \pm 0.8$$

from shrew (3.5gms.) to cattle (700 kg.).

(Source: Altman, Dittmar (1), p. 121.)

A more general argument would consider relations like

$$\left(\frac{\rho AH}{W}\right) = C \left(\frac{\rho H^3}{W}\right)^n$$

ρ = density
W = weight
A = surface area
H = height
C = some constant of proportionality
n = some suitable exponent

to be species specific. The particular form of the law should be characteristic of homologous geometric properties. For example, the classic Dubois data on humans may be represented by

$$\left(\frac{\rho AH}{W}\right) = C\left(\frac{\rho H^3}{W}\right)^{0.575}$$

centered on
H = 160 cm.
W = 50,000 gms.
A = 15,000 cm.^{2}

However, permitting the centered value to stand, it does not abuse the relation greatly to write

$$\frac{\rho AH}{W} = C\left(\frac{\rho H^3}{W}\right)^{1/3}$$

The advantage is that it makes the correlation completely height insensitive. Then the Dubois data become

$$\frac{\rho A}{W} = 11\left(\frac{\rho}{W}\right)^{1/3}$$

Note that this is like the relation given in Altman. We can resolve the difference with the Altman data and take

$$A = (10 \pm 1)\left(\frac{W}{\rho}\right)^{2/3}$$

as a nominal design relation for all mammals.

(b) <u>The 'law' of the development of the aorta</u>.

(1) The aorta substantially extends for the entire length of the trunk, and thus approximately is as long as the trunk. As these relations unfold it will be noticed that we have given the primary development role to the trunk - as the carrier for mesodermic organs. The limbs, as extensions, play secondary roles.

$$L_t = L_a$$

L_t = length of trunk
L_a = length of the aorta (Assumed. Animals exist, with aortas covering the range of lengths from 1 cm. to 20 meters.)

(2) The aorta develops largely with geometric similarity through all mammalian species. The aorta tapers because it tends to preserve the same cross-sectional area (e.g., mean velocity) at every branching level. Thus the aorta may be characterized by its entrance diameter, and by the fact that the length to diameter ratio of the aorta will be approximately constant.

$$\frac{L_a}{D_a} = \left(\frac{L_a}{D_a}\right)_o$$

$$\frac{L_t}{D_a} = \left(\frac{L_a}{D_a}\right)_o = 21$$

(A source of data, for the dog, was Patel et al. (2). Their average aorta was 39.5 cm. length, with 2.0 cm. entrance diameter. Other arteries seem to have length to entrance diameter ratios of about 25 to 1. A nominal value of 21 to 1 was selected for the aorta.)

(3) The average aortic velocity in a resting individual is constant. It is a hypothesis that as an effect of the 'stressful' vibration-acoustic character of the arterial pulse, the growth processes of the cells are so affected as to maintain an approximately constant mean aortic velocity. In fact, the velocity is essentially constant throughout the entire aorta and the main arterial branches. (Their total cross-sectional area remains constant.)

We must assume that this relation holds for the average blood flow in any behavioral epoch. (This period is likely the average over behavioral epochs of the order of 60 days or longer. People tend to change athletic patterns in some such spurts. Females vary activity over the menstrual cycle; athletes often seasonally. It seems only fair to assume with such changes there must be architectural changes to accommodate, but nominally keeping mean velocity constant.)

$$(Q_b)_o = \frac{\pi}{4} D_a^{\,2} \, (V_a)_o$$

$$(V_a)_o \approx 15\text{-}20 \text{ cm./sec. (say 17 cm./sec.)}$$

<u>Derived</u>

$$(Q_b)_o = \frac{\pi}{4}\left(\frac{L_t}{21}\right)^2 (V_a)_o$$

$(Q_b)_o$ = blood flow at rest

(c) The perfusion rate of tissue is proportional (nearly) to the mass (or weight) of the animal. (i.e., it is the weight of the animal which determines the minimum or rest blood flow.)

$$(Q_b)_o = \left(\frac{Q_b}{W}\right)_o W$$

(The literature gives laws from $Q_b \propto W$, e.g., 0.13 l.p.m./kg.; $Q_b \propto W^{2/3}$, e.g., Q_b = 6.5 l.p.m., W = 70 Kg. ; $Q_b \propto W^{3/4}$, e.g., W = 1 kg., Q_b = 0.2 l.p.m., White et al. (3).)

As sources for data on cardiac output and oxygen consumption, Altman (1), White (3), Patterson (4), and Bartels (5) were found. The data covers the range 100 gms. to 2000 kg. (rats to elephants). What is missing for greater certainty is the extension of data to the shrew (5gms.) and to large whales (10,000 kg.)

The data scatter within a range of about a factor of 3 for all species. (That is, if the range is 2 $\pm$ 1, the standard deviation would be about 0.3, or 2 $\pm$ 0.3. Such a scatter of about $\pm$ 15% is not too bad.) There is some evidence that not all species follow one universal power law (within the 3 to 1 band).

As a global design summary, mammals appear to be 'designed' to run with a rest state blood flow near

$$(Q_b)_o = C\, W^{0.9 \pm .05}$$

centered nearly at W = 70 Kg., Q_b = 6.5 l.p.m. (i.e., the human value). The blood flow thus exhibits nearly a unity power law.

(d) Correspondingly, the oxygen consumption and metabolism at rest must also be nearly proportional to the weight of the mammal.

Based on the same four sources used for blood flow, we can find a consistent result.

$$\left(\Delta Q_{O_2}\right)_o = C\, W^{0.9 \pm 0.15}$$

ΔQ_{O_2} = oxygen consumption

centered at about

W = 50 Kg. $\left(\Delta Q_{O_2}\right)_o$ = 0.25 l.p.m.

This is a little higher than the human datum.

The fit for oxygen consumption funnels through a data band that is broader than that for blood flow. Over an appreciable weight range above 10 lbs., correlation is achieved with a 0.9 power.

However this oxygen consumption relation does not seem to hold below 1 lb.; smaller animals seem to 'saturate' in their quiescent oxygen demand, as if such small animals required a minimum of 1-10 cc/min. of oxygen independent of size (for mammals).

This 'saturation' may have one of two causes. It may be a mistaken observation in that these small animals are always 'jittery' - from anxiety - so that they have not been measured in a quiescent state. Alternately they may be 'jittery' as a measure of their higher than normal metabolic activity at all activity levels. The two 'explanations' are not necessarily that far apart.

In any case, the scatter for oxygen data is really much more than for blood flow. This does not provide assurance that either set of data is better than the other. It points up the existence of a measurement problem.

(e) As a test characterization of these results up to this point, we can propose the 'prediction' of an aorta size associated with weight.

If

$$\left(Q_b\right)_o = 6.5 \text{ l.p.m.} \left(\frac{W}{150 \text{ lbs.}}\right)^{0.9}$$

$$= \frac{\pi}{4} D^2 V_o = \frac{\pi}{4} D^2 \left(17 \frac{\text{cm.}}{\text{sec.}}\right)$$

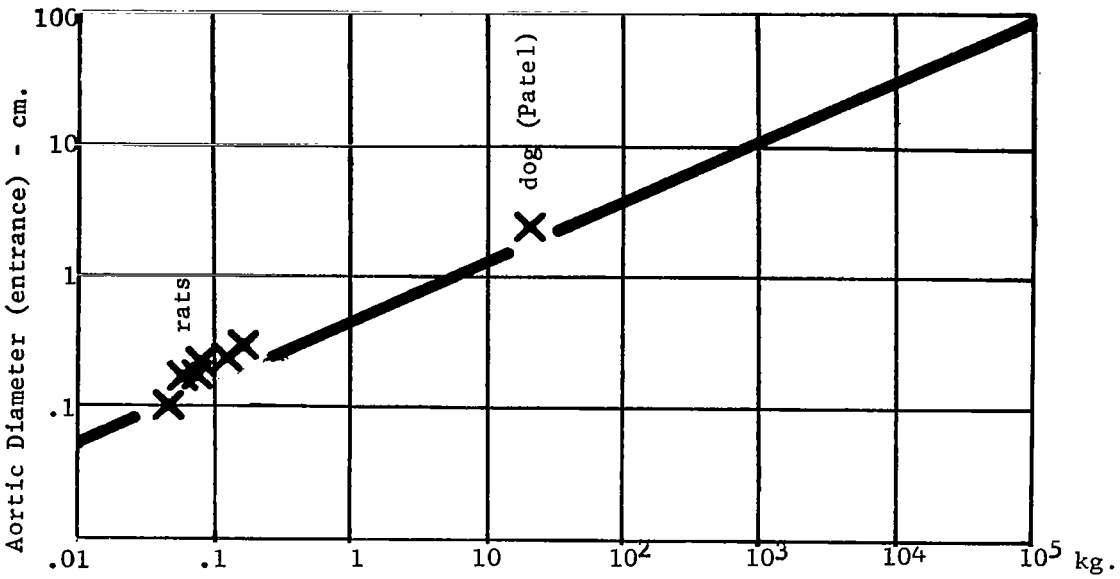
in which we have chosen the value

$$V_o = 17 \text{ cm./sec.},$$

we can obtain the following relation between diameter and weight. (See Fig.).

While speculative, it represents a sharp test of a primary design characteristic of the vascular system. If we can go from weight to blood flow and from blood flow to aorta size then, as we showed before, the design of a space filling vascular system to provide blood flow and oxygen supply can be completed. As we suggested, it is only small 'jittery' mammals that fall out of a unitary series.

In general, we may surmise from studying regional blood flows in various organs and various species, that the specific demand of the large bulk of tissue, tends to cluster within a limited range, when operating at its time averaged state (a range of 4 to 1 in the quiescent state). The average for the whole animal clusters even over a more limited range.



Aortic Diameter vs. Weight for Mammals

Exceptions may be noted in the high specific perfusion of glandular tissue, the kidney and the skin. Of course each organ, when not at its quiescent state, exhibits a large range of change in perfusion.

If the vascular bed is observed at the level of capillaries, a large variation in the number of open capillaries can be noted between quiescence and maximum dilation. The number of open capillaries can vary over about a ten to one range as an average over the entire body. In thermally sensitive skin, capillary dilation is reflected in a flow variation over a 200 to 1 range, whereas in striated muscle dilation results in a variation only over a factor of two in the number of open capillaries.

Representative of the control range of the microvasculature (arterioles and small arteries) the blood flow for the entire body may be varied by a factor of 4-5. Since only a factor of 1½ may be due to pressure increase, the remaining factor of 2-3 is the resistive control range.

If we take the results on quiescent oxygen consumption and blood flow at face value, for which

$$\frac{(Q_b)_o}{6.5 \text{ l.p.m.}} = \left(\frac{W}{150 \text{ lbs.}}\right)^{0.9}$$

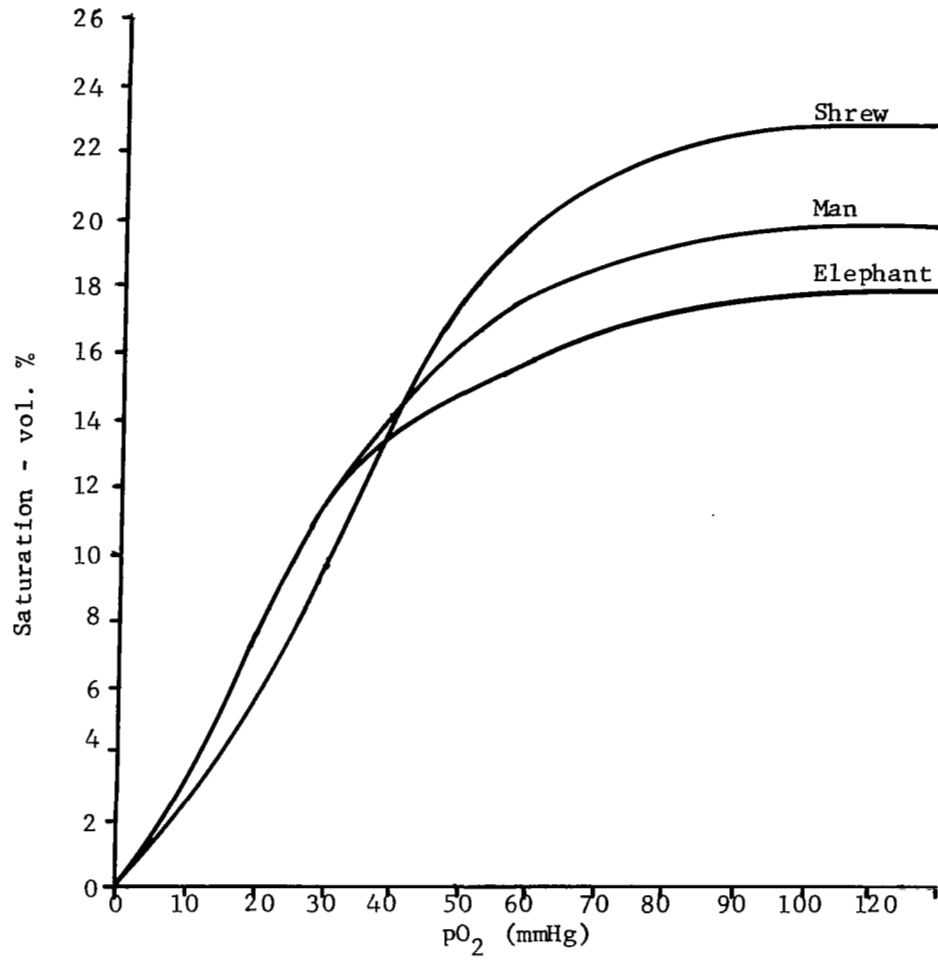
$$\frac{(\Delta Q_{O_2})_o}{0.28 \text{ l.p.m.}} = \left(\frac{W}{100 \text{ lbs.}}\right)^{0.9}$$

we find

$$\frac{\left(\Delta Q_{O_2}\right)_o}{(Q_b)_o} = 5.5\ (\pm 1.5)\ \text{vol. \%}$$

The implied a-v oxygen difference is consistent with human data.

In fact, Bartels provides data on the uptake curve for shrew and elephant, which can be compared with human data.



Normal Oxygen Saturation of Blood
(Source Bartels (5), Ruch and Fulton)

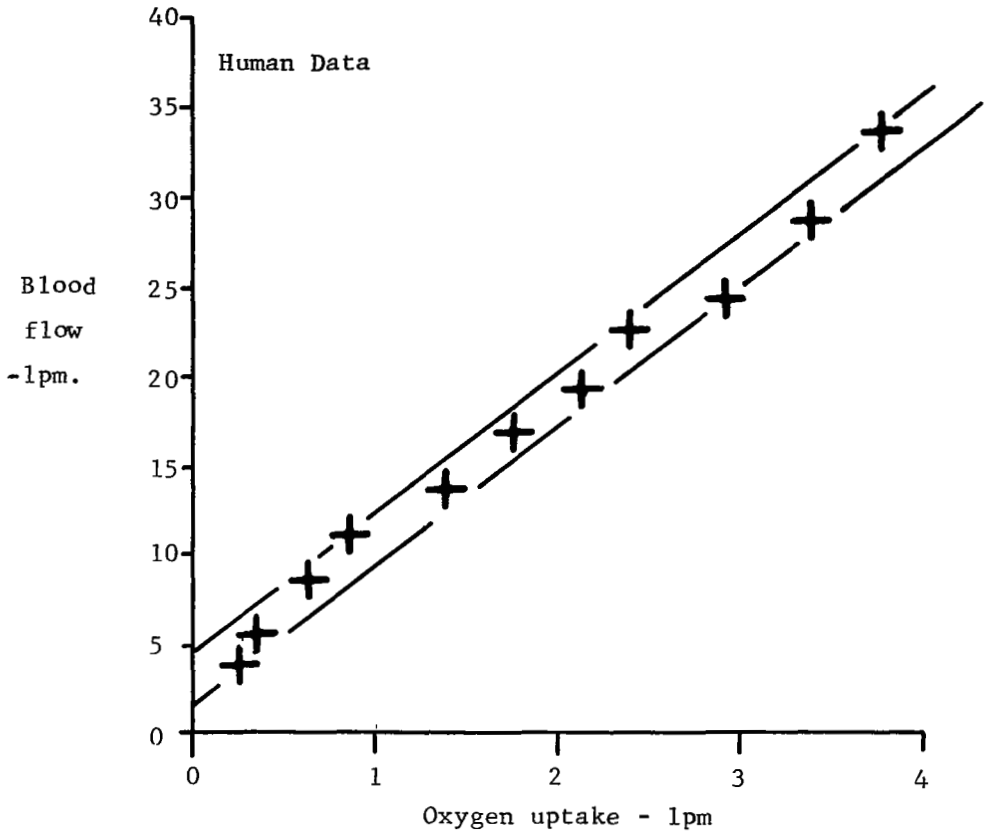
As an approximation these three curves are the same. The human operates with an a-v difference of about 5 vol. % at rest (i.e., between 90 and 40 mm Hg). According to Bartels, elephants operate with an a-v difference of about 5 vol. % at rest (i.e., between 95 and 30 mm Hg). According to Patterson, the cow and giraffe operate between 80 mm Hg and 40 mm Hg. Using the elephant or human curve, the a-v difference would be estimated as 4-6 vol. %. Thus we have consistency in this range.

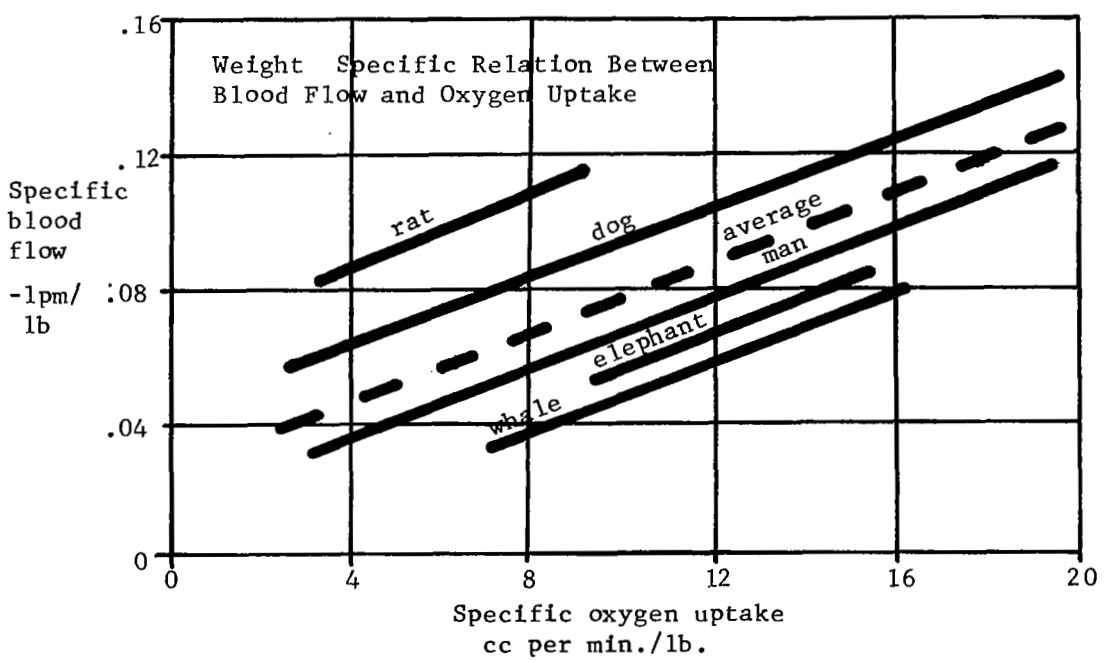
On the other hand, Bartels indicates that the shrew, operating between 90 and 30 mm Hg, has an uptake of about 13 vol. %, i.e., twice as much oxygen. (Also he shows a higher capillary density in muscle for small animals, which tends to be consistent with the observation of increased uptake.) Thus it is likely that the successful small mammals do in fact unload more oxygen at 'rest' than do larger mammals, by a factor of two or so, but they may also have to have a larger specific metabolism in order to survive. The limiting design factor may be an increase in size-specific work to pump the blood in small size animals. The increased capillary density would suggest an increase of 'nutrient' channels.[1]

Thus what we expect in animal design is an uptake that can vary up to 15 vol. % at peak activity, for all mammals, and about 5 vol. % at rest for mammals above 1 lb., rising to 13 vol. % at rest for small animals. As an approximation, the following universal curves for activity can be proposed.

From a number of sources of data on human exercise - mountain climbing, performance in hot climates, performance in the cold, it appears that 3 l.p.m. is the peak sustained aerobic level of oxygen uptake for man. Since it is commonly believed that athletes can exceed this consumption, e.g., consumptions up to 5-6 l.p.m. have been reported, it is noteworthy (Costill (6)) that expert long distance runners could maintain about 2/3 of their peak oxygen uptake of

[1] Our discussion in both the section on the microvasculature and this section on the gross vascular system will be found to be ambiguous on this very critical topic, of what determines the oxygen uptake at the local site. It represents an unresolved conflict in our respective points of view. In one view, following Krogh and Starling, there is a passive diffusion, geometrically determined, from the elementary capillary cylinder. In another view there is an active 'algorithm', depending on the recent past history of usage which determines the operative state of 'tone' of the local capillary bed. In particular, in this view, the 'nutrient' channels are those channels that are smaller than the red cell (e.g., 2-5 μ with 8 μ red cells.). Oxygen supply is obtained by 'stripping' or 'milking' of the red cells that augments the passive diffusion. Reconciliation or extension of such speculations awaits considerably more experimental study. We reserve further comments for the future at such time as investigations have succeeded in providing more experimental data at the level of the functional unit. At this point we are only trying to determine some characteristics that seem to be nearly 'universal' in mammalian tissue.


Human Data
Blood flow -lpm.
40
35
30
25
20
15
10
5
0
0
1
2
3
4
Oxygen uptake - lpm


Weight Specific Relation Between Blood Flow and Oxygen Uptake
Specific blood flow -lpm/ lb
.16
.12
.08
.04
0
rat
dog
average man
elephant
whale
0
4
8
12
16
20
Specific oxygen uptake cc per min./lb.

of 4.7 l.p.m. (i.e., 3 l.p.m.) with a low blood lactic acid level. One may therefore accept 3 l.p.m. as the maximum sustained aerobic level of oxygen uptake.

(f) The operating point. The problem that such physical-numerical analysis of cardiovascular parameters poses is that the implied characterization of the dynamic changing cardiovascular system refers to what the organism can do in a moderate period of time within which a chemical steady state is achieved. The transient response to a herculean weight lift, or to an all out 100 yard dash, or to a terrifying stress was not covered or envisioned in the analysis. However, it also took an entire epoch of development - say 20 years for men (or birth to adolescence in other species) - to bring the system up to its architectural state that it is capable of the sustained performance proposed on a regular basis. Thus we require a careful specification as to the 'operating point' of the system. This requirement is not strange to the dynamic analysis of any other complex system.

Since for the (average) human we may select

$$W = 160 \text{ lbs.}$$

$$(Q_b)_o = 7 \text{ l.p.m.} \qquad \left(\Delta Q_{O_2}\right)_o = 350 \text{ cc./min.}$$

$$(Q_b)_{max.} = 25 \text{ l.p.m.} \qquad \left(\Delta Q_{O_2}\right)_{max.} = 3000 \text{ cc./min.}$$

then

$$(q_b)_o = \frac{Q_b}{W} = 0.044 \text{ l.p.m./lb.} \qquad \left(\Delta q_{O_2}\right)_o = \frac{\Delta Q_{O_2}}{W} = 2.2 \text{ cc./min./lb.}$$

$$(q_b)_{max.} = 0.16 \qquad \left(\Delta q_{O_2}\right)_{max.} = 19$$

are the limits common to most species. Between these limits, the relation is approximately linear.

We may therefore inquire what is the quasi-static or quasi-D.C. operating point of the system? We may surmise that the 'D.C.' or mean operating point is certainly closer to a daily average for biological systems than any shorter period, since they obviously go through a daily or circadian cycle of rest and wake. (It is further clear that the day-averaged operating point is a state that can differ among individuals. The average activity is different for a weightlifter or a marathon runner or steel worker than for a tool room machinist or housewife; than for a business executive or secretary; than for a monk or a convalescent, etc.)

We can try to get some idea of how to view the question of the average operating point by examining variations in weight. Weight variations reflect net caloric balances, and thus net oxygen uptake.

A limited amount of weight data, during a 'normal' life pattern (on a dozen people engaged in normal moderately sedentary to active laboratory activity) showed cycles of approximately 20 days and 60 days. (Example: After sixty days, the differences in weekly 'averaged' weight were 1 lb., 1 lb., 0 lb., 1 lb., 2½ lbs., 5 lbs., 2½ lbs., 1 lb., 1 lb., for a group ranging from 135 to 250 lbs., i.e., the closure - namely the difference in a subject's weight the beginning and end of 60 days - was to an average of 1% of body weight. At 3500 Kcal./lb., the equivalent daily error in closure would represent not much more than 100 Kcal./day, comparable to the 'noise' level in normal eating habits. On the other hand, the weight change range over that period was as large as 7 lbs., 6 lbs., 3 lbs., 4 lbs., 6 lbs., 9 lbs., 8 lbs., 4 lbs. Clearly one could not consider that the body mass and size was stable in anything less than 60 days. The personality types included very rigidly controlled persons, whose weight only showed a 3-4 lb. variation, largely a 3½ day water cycle; and 'widely' fluctuating types who showed 7-9 lb. variation.)

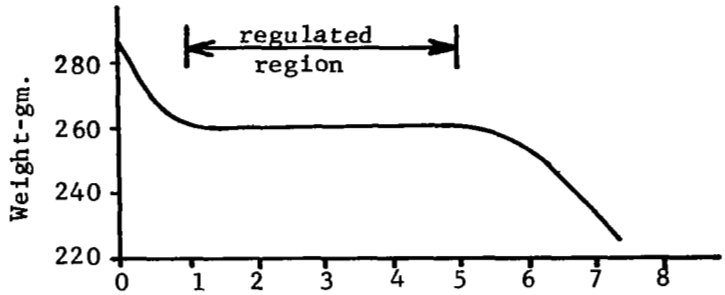
These periods - 20-60 days - represent the fluctuations in 'behavioral' style of living. We suggest that the person adapts to the winds and storms that pour over him. In doing so, the cardiovascular system must also adapt. Thus this epoch represents the growth-adaptive period in the hierarchical regulation of the cardiovascular system.

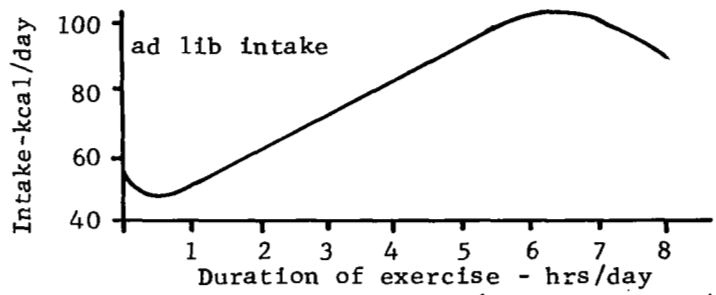
But first we must wash or filter out the daily pattern. The starts and stops of activity and thus change in blood flow over this period are nearly automatic.

What sort of errors, as compared to the longer behavioral cycle, does this leave? As noted in body weight changes, we find maximum derivatives as large as 5 lbs./30 day. (There may occasionally be greater changes - disregarding the water changes - but invariably most people will show such peak changes over a number of days in any extended few months period.) This represents the equivalent of $\pm$ 600 Kcal./day activity. At levels of daily intake of 2000 Kcal./day, 'errors' of 600 Kcal./day are quite appreciable. Thus one may not a priori guarantee a (dynamic) steady state balance in oxygen uptake and thus in blood flow in the cardiovascular system in less than 60 days. The residual errors for longer periods are then as large as minor fluctuating 'emotional' or other noise errors.

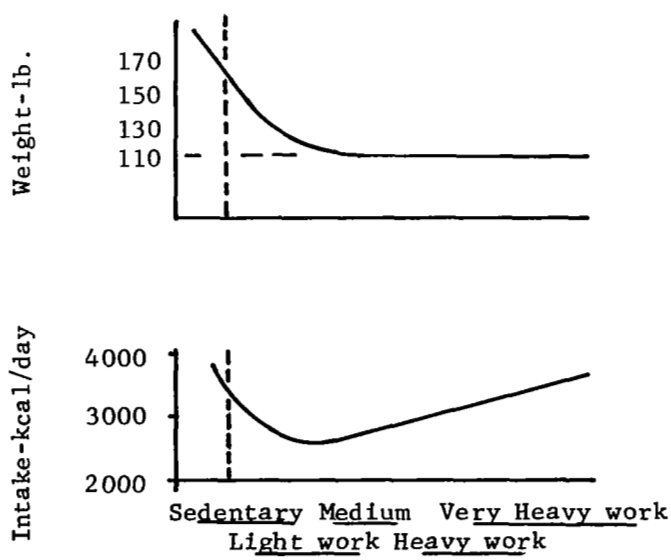
Where does one operate on this longer time scale? Is there system regulation? Fortunately, we have located an answer in some work which has been inadequately highlighted. The work predominantly originates from Mayer (7); also see Parks (8).

In rats, and also in man, the following steady state response was found.





Steady State Response in Rats (Mayer, Thomas (7)).



Steady State Response in Man (Parks (8)).

These data are very elegant and clear data on true regulation in a biological system. Note there is a long term regulation of weight - if the animal is forced ('self' forced or forced by the environment) to be active. Food is proportioned to maintain that weight. The regulation breaks down at low activity. The system 'gain' becomes indeterminate at that low activity. On the other hand at too high activity, the animal 'fatigues' or debilitates. This confirms what we have found in thermoregulation studies, summarized as "The human is designed for moderate activity". The detail we missed was that there is a broad regulating zone. We determined, by adding increments of activity to normally sedentary people, the approximate measure that the human is 'designed' for. It is a minimum of about 2 hours of normal walking activity per day added to sedentary activity, or 40 minutes of swimming. (These results were hard won from a variety of human experiments.) However, the Mayer data portray the results much more broadly. What the data do not provide are the time constants associated with reaching the regulated state for systems that are outside the regulated range.

We submit that this activity range fits all of our comments with compatibility; namely, the human is designed for the equivalent of 2-10 hours of added walking activity per day, or 40-200 minutes of a heavier activity like swimming. Brouha's studies on daily work tasks suggest information about the optimal scheduling of such activity. Since such scheduling is governed by parameters which represent blood flow, then the average blood flow is fixed for the individual in that temporal epoch. Correspondingly the average metabolism for that activity is tied down.

However, the universal oxygen uptake - blood flow curves must hold. This does not mean solely short term compatibility with the dissociation curve. (One need only speculate about his operating condition the first time the amateur tries to swim three hours, or run for two hours. Or one can refer to the issue of the Washington Post in which a brave Pierre Salinger took off at the behest of his leader, John Kennedy, for the opening of the 50-mile hike program. Recollection has it that he quit after 6 miles.) In short term, unless the activity is extremely strenuous, most human systems can hang on to an activity for two minutes; or if less strenuous for 7 minutes. (You cannot get an out-of-condition person to swim more than 2 or 3 laps; or to high jump, etc.) To achieve a given operating state, you must take the time of a training period to build up what can be built up at the existing operating point of the system, or of a period of degradation to otherwise modify the operating point. If the person is hundreds of pounds overweight, it may take a long heroic effort to train him; if only moderately 'out of shape', it becomes a matter of weeks to months; namely, the period of 20-60 days is not inappropriate.

What has to happen is an adaptive growth of muscle tissue (or unfortunately of fat tissue); a readaption of the water depots (namely, to a significant extent one might view the body as water and fat supply depots for the working system); and an adaption in the vascular system so as to operate on the universal curve. This means that the adaptive growth algorithm includes means for moving both the blood flow availability and the oxygen availability toward a correlated operating point. Speculatively, we can suggest that the

effect of exercise is to decrease the number of capillary shunts; that there is considerable significance to heat and temperature signaling in the growth and regulation of the state of the capillary beds; that there is considerable significance to be attached to stretch or vibration signals relative to shifts in the arteriolar control level, or growth in the arterial tree.

Above all, we believe the case for system instability and indeterminate gain at zero frequency in the cardiovascular system is beyond question, and the case for hierarchical regulation of its state is also beyond doubt.

With regard to the CV system, what this means is:

You cannot stand still for very long. While standing, the system is designed to operate with a degree of mobility almost all the time. If one 'freezes' the posture, it is the venous system which is very soon affected.

You cannot lie still for long. (Here we refer to the case of days of bed rest.)

Consider what is affected in case of prolonged bed rest. We may estimate the following sequence of effects,

fatigue
lassitude
irritability
disorientation
boredom
irregular sleep
constipation
digestion difficulty
sores
water distribution disturbances
muscle atrophy
catastrophic pooling of blood in dependent veins.

While a 'sick' or debilitated person may be willing to accept the enforced state, an active healthy person can hardly tolerate it. One must note how many of the effects are related to the cardiovascular system.

Thus the response of the cardiovascular system, even of its 'high' frequency beat by beat performance, is related, by system overlay upon system overlay, to the active 'low' frequency behavior and cannot be successfully divorced from it.

Therefore, as we have attempted to stress, physiological equilibrium does not come into existence until one or more major behavioral equilibria are achieved. In the case of the CV system, this refers to the 'episode' of life which contains man's activity patterns.

The interesting aspect is that even though one thinks one makes minute to minute decisions on food, activity, water, etc., there really is an integrated overall level developing or being maintained which one's apparent conscious decisions only subserve.[1] Thus the better question one must ask in examining the CV system is not what did one do yesterday, but last season.

Obviously these involve very subtle and deep seated mechanisms. While the conventional wisdom is that the central CV control comes from the CNS, in particular the autonomic system as a high frequency system, we suggest, quite speculatively, that the rhythms and tides arise from the wash of the endocrine systems, in which the autonomic system is manipulated as a low frequency system.

(g) The arterial 'pulse'. Of course the CV system is energized, beat by beat, by the ejection pulse of the heart. As an approximation, a nearly autonomous character of myocardial tissue associated with the species determines the intrinsic heart beat. Thus rest flow is nominally related to the beat period by

$$\Delta\tau_o = \frac{\Delta V}{(Q_b)_o}$$

$\Delta\tau_o$ = basic heart beat (characteristic of the shortest period autonomous tissue in the heart - usually the SA node)

ΔV = stroke volume of the heart

$(Q_b)_o$ = basic flow rate

This basic heart beat is approximately associated with animal size, as the following graph (from Altman (1)) shows.

However, then there are sympathetic and parasympathetic nervous changes in heart rate, which can be expressed as

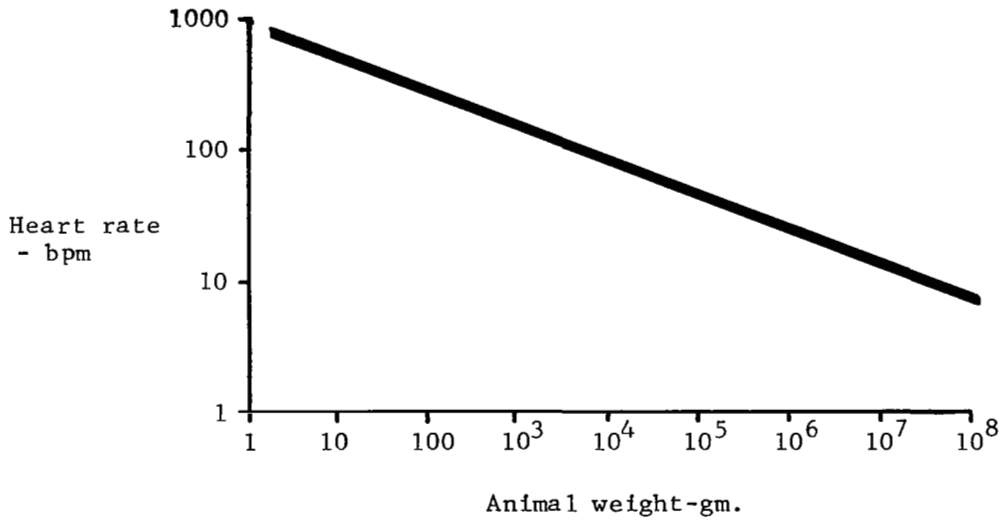
$$\Delta\tau_o + \Delta\tau_p - \Delta\tau_s = \frac{\Delta V}{Q_b}$$

[1] To get this out of the realm of speculation and more into the realm of 'hard' experimental data, we can illustrate with two recent references. On one hand, Terman and Terman (9) illustrate a remarkably uniform large amplitude 'volitional' behavioral rhythm at circadian time scale in rats; on the other hand Wolf (10) exhibits an inhibition of an autonomic function, "an extraordinary restraint of cardiac reactivity," that was shown to develop in a woman over a period of days. On the day on which an experimental stress was to be focused, she clamped her pulse at a remarkably constant rate.

$\Delta\tau_p$ = 'parasympathetic' effect (generally a slowing down)

$\Delta\tau_s$ = 'sympathetic' effect (generally a speeding up)

To some limited extent, in 'closed loop' operation, there is some minor changes possible in stroke volume from sympathetic effects. In 'open loop' operation, the changes can be large. Thus mean blood flow occurs in association with changes predominantly in heart rate, and to a lesser extent in stroke volume.

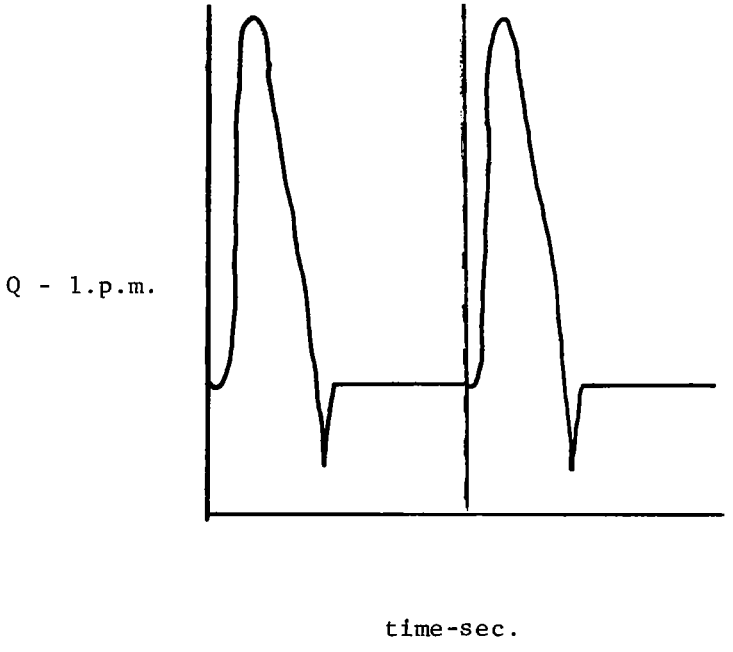


As a result of extensive training, the heart rate, for the same task, is diminished. Part is due to greater efficiency of the task, but more pointedly, the resting heart rate is reduced. This is the resultant change in development that permits higher peak demands.

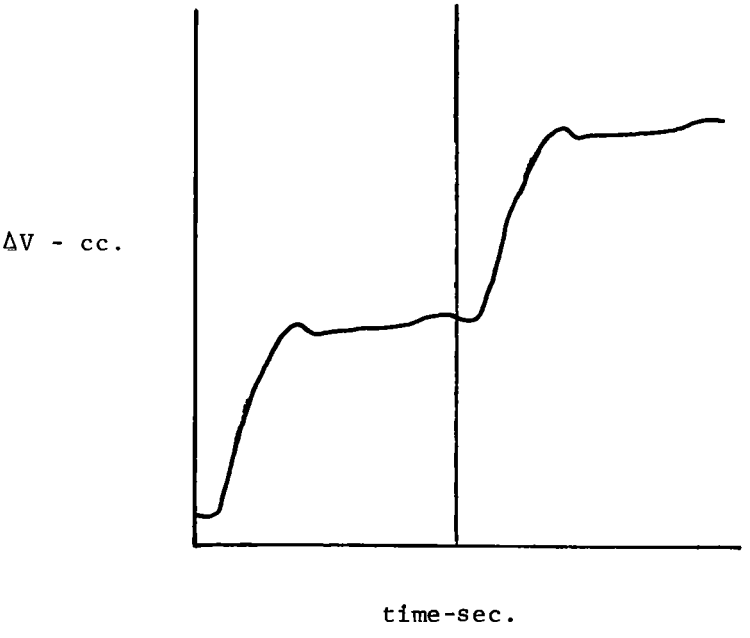
As far as the high frequency arterial pulse is concerned, the approximate characteristic of the CV system is described by the 'windkessel' characteristic of Frank (1899).

To illustrate:

A pulse of flow ejected by the heart

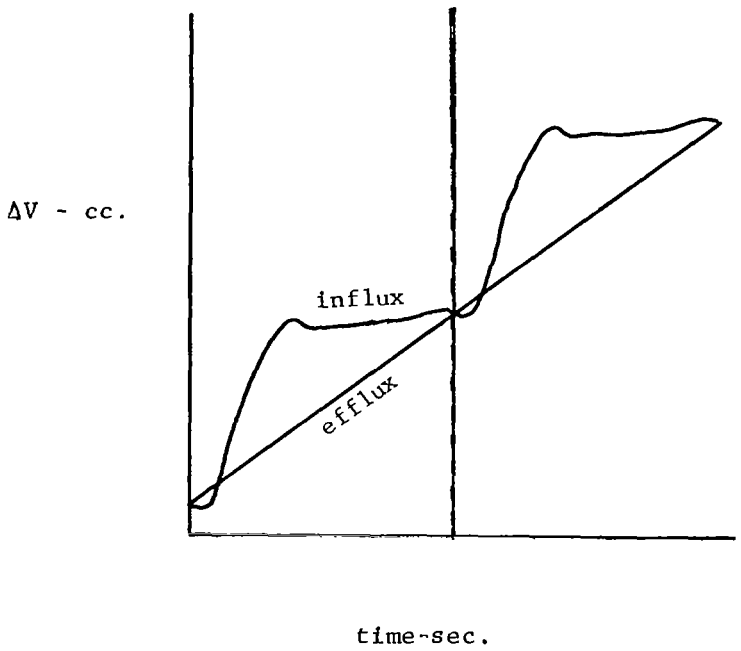


may be integrated, to exhibit the step of stroke volume that has been injected by the heart into the aorta

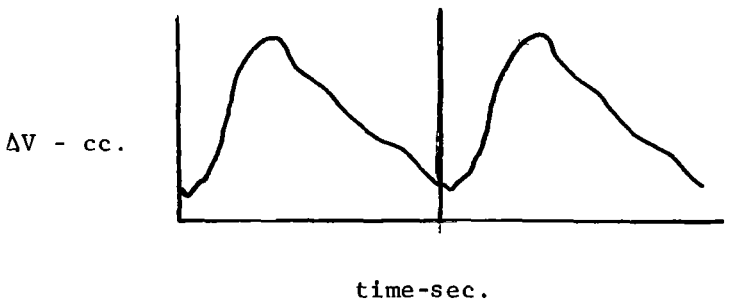


(since the pulse has a finite width, the 'step' is actually a 'ramp' of volume).

However, actually there is a efflux of volume - at approximately a constant rate because of the high pressure pump characteristic of the heart through the terminal resistive bed (essentially associated with tubes less than 300 μ in diameter). Because of the approximate 25 to 1 ratio of length to diameter, the arterial lengths associated with the resistive layer is less than about 1 cm. in length.

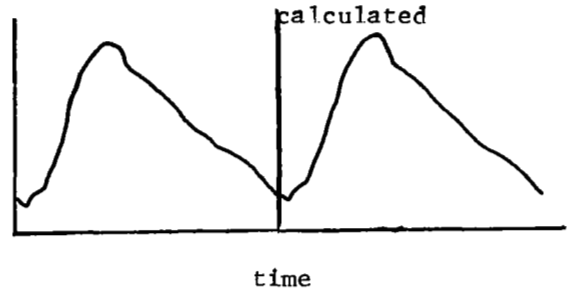


The net volume difference, stored in the elastic aorta and branchings, forms a triangular wave of volume

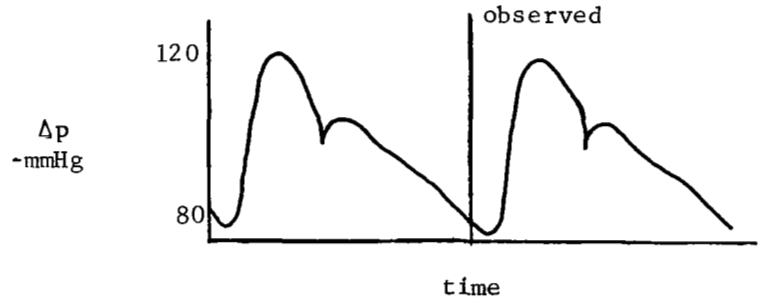


The nearly constant elastance (over the operating range) transformed the triangular wave of volume into a wave of pressure

$$\Delta p = \left(\frac{dp}{dV}\right) \Delta V$$



When compared with an actual pressure wave, except for some high frequency detail, they are quite similar.



Thus the essential character of the arterial pulse is explained by the resistive characteristics of the microvasculature (which is fixed for any short period of time), and the elastic characteristics of the arterial tree. The net effect is that the diastolic-systolic difference in pressure is not an independent variable but depends on the stroke volume and elastic characteristics of the vascular tree.

With this introduction, at the organ level, we now propose to examine the cardiovascular system, as one demand subsystem of the metabolic system. The cardiovascular system serves a primary function of distributing oxygen, where the term 'primary' signifies that the substance in blood whose supply rate by arterial blood comes closest to the tissue consumption rate is oxygen. The

many other substances convected by blood to tissues can be regarded collectively under three headings: fuel supply, water supply and information flux.

2. Temporal Fractioning of Cardiovascular - Metabolic Phenomena

A useful prelude to this discussion, in keeping with the general themes that have been touched upon, is a temporal classification of physiological functions for which a metabolic and an associated cardiovascular substratum must be found. The metabolic events will serve as a background against which to evaluate cardiovascular phenomena. The times chosen ought to be regarded in the following light. Suppose one wishes to graph a physiological phenomenon unfolding in time. The ordinate will be subdivided into appropriate units-the metric of the process under study. The abscissa will be subdivided into appropriate temporal units: growth of a human in years or decades, growth of a rat in months, growth of a bacterial population in culture in hours, growth (and decay) of an action potential in fractions of a millisecond, and so forth; yet the linear dimensions of these graphs would be the same. So each time specified below might be regarded as the temporal unit of the major divisions of an abscissal scale.

(a) 0.1-0.25 sec. - Individual motor units release energy within this time scale. It may also be taken as the shortest interval over which a behavioral change could be detected at the organismal level (e.g., the change from a relaxed to an alerted posture, a strike of a piano key, a change in facial expression, and so forth).

(b) 0.5-4 sec. - Cyclic events of blood pumping and breathing. It is also an approximate transit time for single red cells through a capillary path.

(c) 10-25 sec. - Major postural changes occur in this time scale. It is also an approximate time constant for change in the oxygen stores of the body when oxygen consumption is in the upper physiological range. It is within this epoch that continuity of flow within the cardiovascular system is seen (e.g., left and right ventricular outputs are equal and cardiac output equals venous return.) It is the approximate duration of a maximal external power development (e.g., 200 yard dash, which is an all-out dead run usually with breath held).

As an example, Sargent (11) shows data for a subject well-trained both at sprinting and participation in physiological measurements. The maximum oxygen consumption that the subject could sustain was approximately 3.5 lpm. His maximum non-steady state rate of oxygen consumption was about 35 lpm, and his maximum oxygen debt was about 14 liters. These figures extrapolate to an energy burst of about 25 seconds (about a 250 yard dash).

(d) 100 sec. - An approximate time constant governing the increase in oxygen intake rate in all-out exercise (11) or increased activity (12). Metabolic power decays with this time constant upon cessation of a task. Metabolic power averaged over the entire organism fluctuates with this appropriate period (13). Red cell flow through small single capillaries also fluctuates with

approximately this period (14). There is evidence (15) for oscillations of fuel concentrations in blood with this period. It is a plausible approximation of the time over which oxygen balance will occur within functional units in microscopic regions.

(e) 400 sec. - This is an approximation of the time period over which oxygen balance must occur for the whole organism under non-exertional work loads. It can be regarded also as an approximate settling time for cardiac output after changes in activity (12).

Benzinger's measurements (16) of oxygen consumption and tympanic membrane temperature show correlated fluctuations which have a period of approximately 400 sec. These and other data (6, 12) suggest that 400 seconds is an approximate settling time for the apportioning of cardiac output between cutaneous and other major systemic circuits. A striking example is the approximately 7 minute period in skin temperature in resting human subjects (13, 17), indicating significant variation in cutaneous blood flow at this period. Considering the improvement in the technology of blood flow measurement in unanesthetized animals, there is a notable paucity of temporal data on the apportioning of cardiac output among the major beds in undisturbed experimental animals.

(f) 1200 sec. - The approximate time constant for CO_2 stores in the body fluids. Probably as a concomitant is the fact that there is a prominent oscillation in ventilation with this approximate period (13, 18). It is also the time constant for recovery of oxygen consumption and CO_2 production from maximal rates to steady state values of mild exercise, after a 30 second burst of all-out exercise (19). Behaviorally, it is likely that spontaneous shifts from a fixed behavioral focus on a single behavioral modality take place in this time scale. If true, then shifts in the blood distribution have to take place in this time scale, also.

Examples of singularly focussed activities for which 20 minutes is an approximate epoch are sitting still at a lecture, intense sexual involvement, eating to satiation, the twenty minute dramatic content of a half-hour radio or television program, one act in a play, and so forth.

(g) 90 min. - Evidence is accumulating for a 90 minute epoch in behavioral activity (20). While first discovered in penile erection and then as REM epochs in sleep, it has been detected within waking performance.

(h) 3½ hrs. - In this epoch, thermal balance occurs. Recent evidence has been put forth for the involvement of cortisol at this scale (21,22). The shift between 'volitional' behavioral modalities often involving extensive change in motor activity, occurs approximately on this time scale. At this epoch in time, the cortex, appraised of the status of many organ systems, commits the organism to a long term task. A typical example is the commitment to a morning, or afternoon, or evening work session; going to the theater, reading a book, playing bridge, etc. The operating points of organ constellations change.

(i) 24 hours - There are numerous circadian rhythms not only in the chemical constituency of body fluids, but, quite obviously, in behavior as well, (e.g., the periodic nature of episodes of the day, first-wake, feed, eliminate, motor activity, etc.) A striking illustration in behavior has been recently reported (9).

(j) 3 1/2- 7 days - Within this epoch water and electrolyte balance is achieved. There is also a behavioral-social cycle at this period, out of whose cues the week arises as a temporal unit.

(k) 20-60 days - Some major behavioral rhythms occur on this time scale. The most certain of these is the menstrual cycle. Twenty to sixty days is an approximate settling time for the constellation of responses which make up athletic training - changes in vascularity, in the resting stroke volume-heart rate relation, and in muscle mass. Adaptive changes of a behavioral nature - anxiety, creative effort, change of pace, etc.-take place in this time scale. Major activity pattern changes may be noted.

(l) One year - Seasonal cues operate, leading to rigidly deterministic behavioral cues in lower animals and more subtle influences in man, such as differences in 20-60 day average activity.

(m) 5-10 years - In man this period might be regarded as constituting a mini-career. The "seven year itch" applies not only to one's choice of partner but to one's choice of major focus in daily activity. Ten years is a significant career segment.

(n) 20 years - A major life epoch associated with the transition from dependency on parents, to being a parent, to dissociating from one's young and assuming the responsibility for one's parents, to decline and re-establishment of dependency on others.

(o) 50 years - Within a couple of generations appreciable change in body size, with dietary changes, have been noted in this century.

Metabolic and regulatory concomitants play out in each of these time scales. Current knowledge, derived largely from acute experiments, focuses on those mechanisms which are involved in the first 6 or at most 9 epochs listed. The same may be said of CV phenomena.

3. Rationale for Focussing Initial Discussion on the 400 Second Region of the C V Spectrum

It is usually intuitively customary to begin analysis of the cardiovascular system with the 400 sec. level. It may be regarded as the first integrative level in the sense that a variety of cardiovascular mechanisms operating on this or shorter time scales interact to effect body-wide oxygen balance and to produce an approximate 400 sec. level. It may be regarded as the first integrative level in the sense that a variety of cardiovascular mechanisms operating on this or shorter time scales interact to effect body-wide oxygen balance and to produce

an approximate 400 sec. settling time in cardiac output and regional flow adjustments. At this time scale, the regulation of cardiac output (Q_b)[1] can be regarded as the regulation of oxygen convection ($Q_b C_a$), where Q_b = cardiac output and C_a is the concentration of oxygen in arterial blood.

It is desirable that we characterize our 400 second fragments. This time scale is too coarse to see the processes that make up the individual heart beats. Yet we must recognize that pressure and flow are obtained from an intermittent pump, beating at a near constant frequency. It is the beat to beat variation that we are neglecting. Similarly, we are neglecting the epoch by epoch regulation by the baroreceptor reflexes. Its cardiac action tends to be reflected in changes, through the vagus, at 5-15 second epochs (23) and there is a 20-60 second time constant in its actions on resistive vessels (24).

Thus 400 seconds emerges as a near steady state fragment of life in which both the oxygen uptake and blood flow are in accordance with the Fick principle throughout the system (i.e., no changes in volume storage of oxygen). Also, at the 400 second time scale, anaerobic transients are neglected, and the oxygen supply rate via the arterial stream ($Q_b C_a$) is a sustainable excess (>1.3) over total body oxygen consumption.

Now there are many body states possible that can be demonstrated and maintained at near steady state. We no longer are talking about rapid changes, or bursts, but modalities of behavior. In a behavioral modality, it is a complex constellation of organs and systems that are operated compatibly at a specific operating point. For example, it is difficult to make use of the leg muscles at a high exertion level and perform an intensive mental task; it is certainly more difficult to add a third task, such as eating simultaneously.

Consider for a moment some of these prolonged and sustainable task states. One can rest, sleep, walk or run at moderate speeds, attend intensely in a perceptual mode, etc. Such states are marked by the fact that near steady state blood flow distributions can be achieved. The time averaged metabolism and blood flow over a number of 400 second intervals is not very dissimilar from the time averaged metabolism over any one 400 seconds, as an irreducible or atomistic temporal fragment of 'normal' total body metabolism.

It is useful, therefore, to start the analysis on the values of some of the key system variables of the cardiovascular system for such 400 second epochs over the range of rest to sustainable exercise.

We should note in passing that exercise is only one of perhaps twenty behavioral modalities. The cardiovascular concomitants of exercise are prominent, sustainable, and obviously important, so exercise is a logical modality to consider in the present context. The cardiovascular concomitants of other behavioral modalities are too poorly defined to make their exploration profitable.

[1] In accordance with engineering custom, we will use V to denote volume, Q to denote volume flow.

(The cardiovascular cycle associated with a single feeding is aperiodic, and it is difficult to associate a steady state description of cardiovascular events unless one treats a large number of feedings. The scale for the feeding steady state far exceeds 400 seconds).

However, we must also think of the system averaged over even larger periods. 'Short' term adaptive changes in the heart and vasculature likely extend out to 20-60 days. Thus, it is really 3 epochs that must be considered for the cardiovascular system in an integrative view. The first is the 400 second epoch; the second is the daily epoch of life over which the duty cycles of different atomistic activities occur; and the third is the 20-60 day epoch, in which the system clamps on a style of life with an approximately singular behavioral outlook (whether to loaf, work intensively, garden, brood, exercise every day, etc. In many animals, the 'behavioral' outlook is often governed by seasonal changes).

We must highlight the daily cycle because the transition from 400 seconds to the day has passed over two intermediate periods which have to be studied if fairly high accuracy is to be achieved in the system's description. For segments less than a 1200 second epoch the Fick principle for CO_2 cannot be used with high accuracy throughout the entire body (chemical equilibrium is not insured). For segments less than 3 1/2 hours heat balance cannot be assured with high accuracy throughout the system (thermodynamic equilibrium is not insured).

The essential point being made is that one is always threatened, in evaluating average system parameters, by the potential for error in the neglect of storage terms in balance equations.

We only anticipate modest coupling between the achievement of CO_2 and thermal balance and prominent cardiovascular concomitants. In other words, the cardiovascular spectrum is relatively less interesting to us in the period range 400 seconds to 24 hours, and 24 hours to 20-60 days.

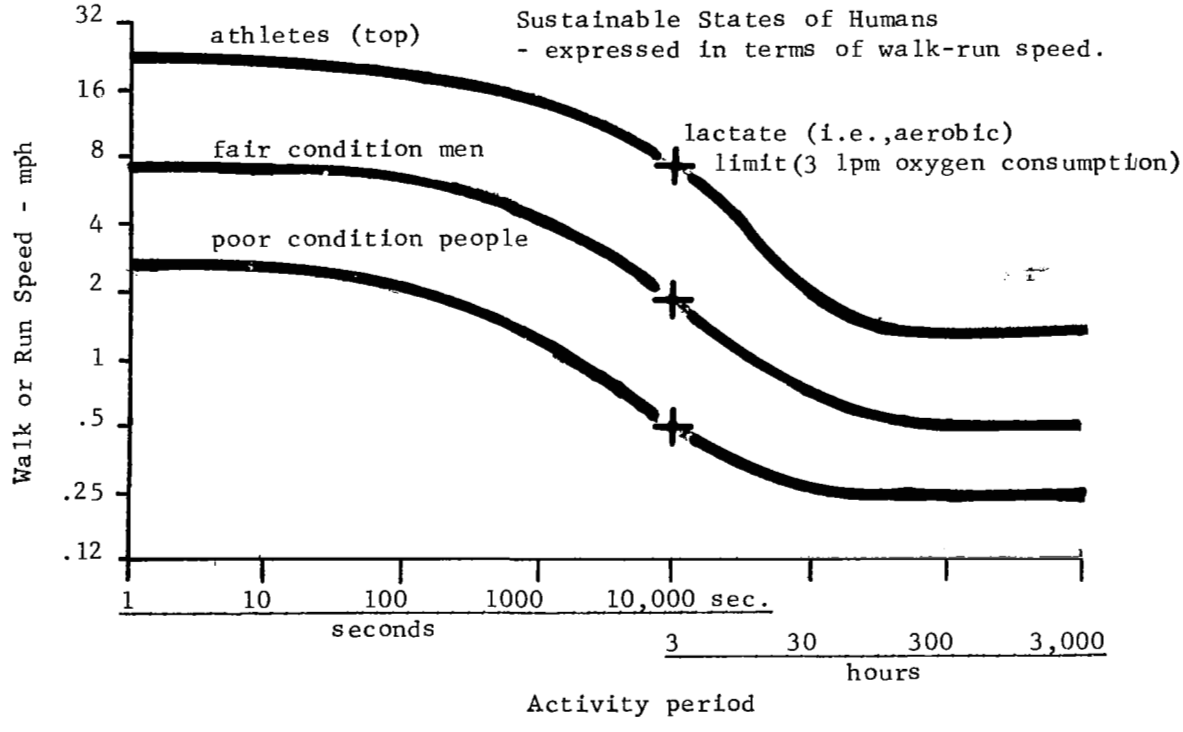
Our concern is with beginning to characterize the system's performance in some deterministic fashion. It will be found, that the characterization is confused, in a number of ways to study the system , by the effects of body posture, body size, exercise, adaptation, and by age and sex. We can tease these effects apart, to some extent, by the following observations.

(a) To some extent, the effects of posture can be stated. Cardiac output tends to be somewhat greater at rest in a supine position than at rest in a sitting position, and, in turn, it is greater than at rest in a quietly standing position. It is difficult to characterize the standing 'rest' position in man as a ground state, because minimal motions prevent hydrostatically dependent venous pooling, which has large inverse effects on cardiac output. Thus 'rest' generally refers to a state of minimal, but normal and real motion in which venous pooling is prevented.

(b) Roughly, the cardiac output is proportional to body weight; the effect of past exercise has little or no effect on the nominal cardiac output at rest. The main effect observed is that cardiac output, of course, increases with exercise. The most specific measure of the exercise state is oxygen consumption.

One would therefore like to think of the cardiovascular response as sampled from an epoch of an individual's life in which his activity pattern can be characterized. The adaptation period is minimally of the order of months (often seasonal), and often tied to the emotional-career-medical status of the system. Within such a characterization, we should really examine the cardiovascular response arranged over the day; for it is predominantly the day which represents the cardiovascular duty cycle. (Idealized as 8 hours of 'work', 8 hours of sleep, 1 hour of eating, 7 hours of 'recreational' activities). However, if one examines the metabolic rate of the system (1200-3600 cal/day), it is clear that the daily average of metabolic rate is nearer rest than near its maximum oxygen consumption (i.e., nearer 1/4 lpm oxygen than peaks at 3 lpm oxygen).[1]

[1] A somewhat speculative, but apt, summary of the sustainable mean or steady state conditions of a man is contained in the following figure. It presents the average activity load of a man, expressed as a running or walking velocity versus the time over which that activity load can be maintained. What is also indicated are the limits at which aerobic activity can be maintained. Curves are developed for three categories of man - athletes in peak condition, men in fair condition, and people in poor condition.



Thus we may think of a posture-independent cardiovascular performance taken at a variety of 'steady state' activity levels - minimally 20-30 minutes at each level, for a crude characterization; or maximally hold for 3-4 hours at each level for more accurate characterization. The duty cycle for the day can then be made up from a patterned sequence of such segments. The practical difficulty of running 24 hour experiments leads one to such sampling. The available experimental evidence in activity has been usually gathered from subjects doing sustainable exercise on a treadmill or bicycle ergometer. It is within the time range of some such maintained steady state activity that we are attempting to view 400 second fragments of CV performance.

4. Some Elementary Parameters in the Cardiovascular System.

Because of the intermittent nature of the heart pump, the cardiac output is described by the kinematic relation.

$$Q_b = f\Delta V$$

Q_b = blood flow from the heart to the arterial system (cardiac output - cc/min.)

f = heart rate (beats per minute)

V = stroke volume (cc per beat)

or alternately, by the reciprocal relationship

$$\Delta V = \tau Q_b$$

The Fick principle - a kinematic relation - in steady state gives

$$Q_b \Delta C = \Delta Q_{O_2}$$

ΔQ_{O_2} = oxygen consumption (cc/min)

ΔC = change in oxygen concentration between nearly saturated arterial blood and mixed venous blood (cc O_2 per cc blood)

Because of the chemical nature of the metabolic process, equivalent to an oxidation of a limited range of fuels, as an approximation, the average heat of combustion is nearly constant, so that

$$M = 4.8\ \Delta Q_{O_2}$$

ΔQ_{O_2} = oxygen consumption (lpm)

M = metabolism (Kcal/min)

4.8 = average heat of combustion (Kcal/l O_2)

Because of the near constant nature of the arterial pressure arising from a pulsatile source (but feeding into a large elastic capacitance), a nominal peripheral resistance can be defined.

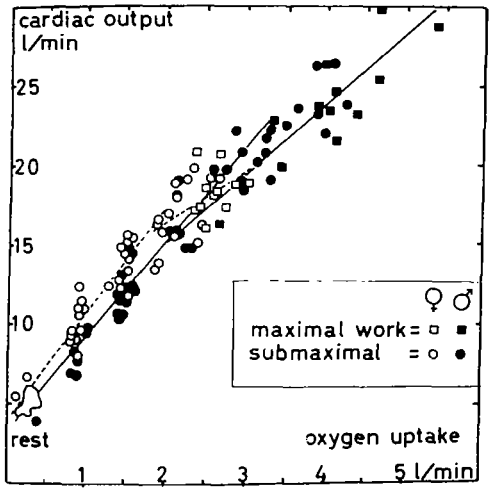
$$R = \frac{\Delta p}{Q_b}$$

Δp = pressure drop from arterial to venous system (mmHg)

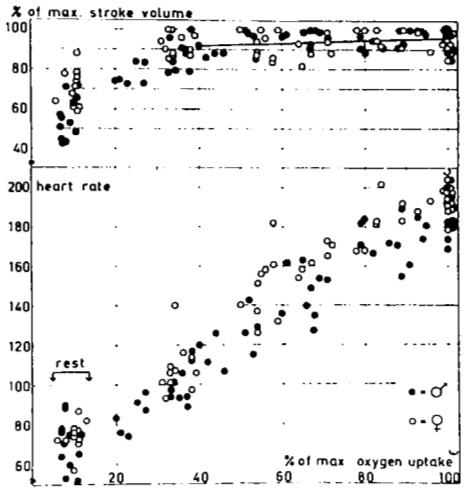
R = peripheral resistance (mmHg/lpm/min.)

5. The Range of Values of the Elementary Parameters in Rest and Exercise.

The following figures from Astrand et al. (25) illustrate exercise-dependent changes in cardiac output, stroke volume, heart rate, and arterio-venous oxygen concentration difference. The studies were performed on a bicycle ergometer. The rest values of stroke volume and cardiac output deserve comment. The earlier work of Bevegard, Holmgren and Jonsson (26) showed differences in these variables between the quiet sitting and quiet supine postures. The differences are explicable in terms of the adverse effect of pooling of blood in dependent veins on cardiac output. With movement, however, even at very light work loads, stroke volume and cardiac output increase, so that there is an almost discontinuous jump in the two variables with the first increment in oxygen consumption above the rest value. With further increases in work load and oxygen consumption cardiac output rises continuously but stroke volume does so only slightly.

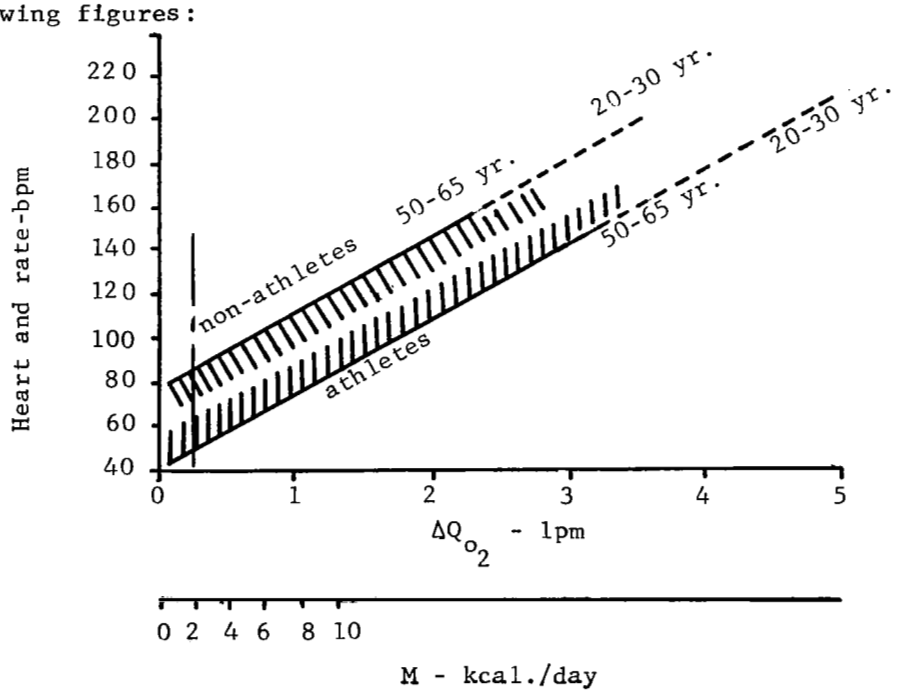


Cardiac output versus oxygen consumption at different work rates in male and female subjects (25).

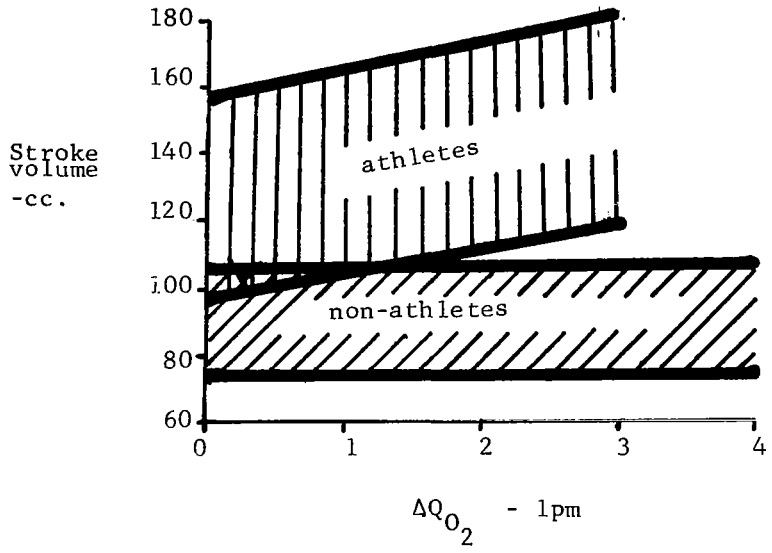


Heart rates (lower) and relative stroke volumes (as % of each individual's value at his maximum oxygen consumption) versus relative oxygen consumption in the same group of subjects as before (25).

A broader summary of a considerable amount of literature is shown in the following figures:



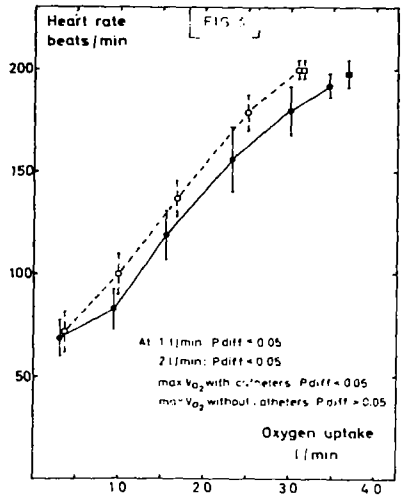
Heart Rate vs. Relative Oxygen Consumption



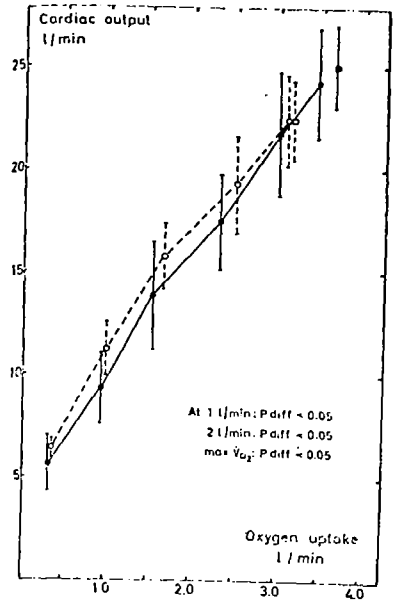
Stroke Volume vs. Oxygen Consumption

Across a 4 month epoch of physical training (cross-country running), Ekblom, et al. (27) found the following changes within a group of 8 subjects at given rates of oxygen consumption (see the following figures) - decreased heart rate (though not at rest), decreased cardiac output (though not at rest), and increased arterio-venous oxygen concentration difference. Oxygen consumption - specific stroke volume was unaltered except at the maximum rate of oxygen consumption, where the value after training exceeded that before. It is interesting to note that work load-specific oxygen consumption changed little after training except possibly in the mid-range (900 kilopound-meter/min). A subsequently reported study by Douglas and Becklake (28) on four hockey players immediately before their seasonal hockey training commenced, and after four months of training, showed similar findings, although the work-load specific heart rates were more greatly reduced after training than in the Ekblom study, and the work-load specific oxygen consumption was not observed to be altered. (It should be noted that in neither study did the subjects train with the bicycle ergometer on which the measurements were made. Thus, the question of whether specific training can appreciably improve efficiency was not directly answered by either study.)

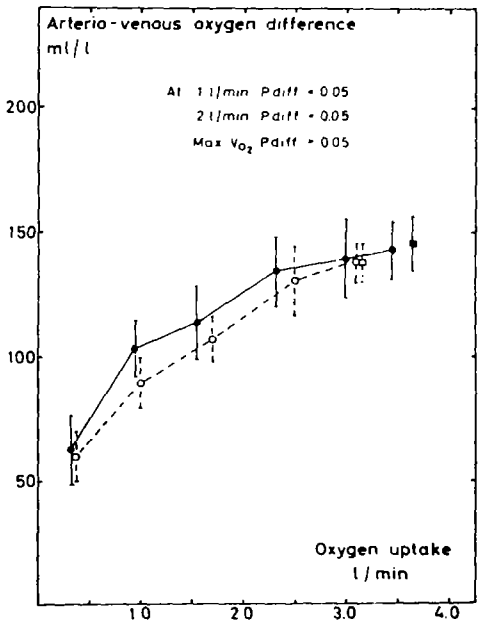
From these figures, we can conclude that with training, the cardiovascular system undergoes 'adaptive' changes which enable it to function at lower oxygen-consumption specific heart rates, cardiac outputs, and mixed venous oxygen concentrations. The last-named adaptation signifies an altered distribution of cardiac output in favor of exercising skeletal muscle, because the observed concentrations of oxygen in arterial blood were unchanged. At rest and at peak aerobically maintainable blood flow, however, cardiovascular functioning was little altered.



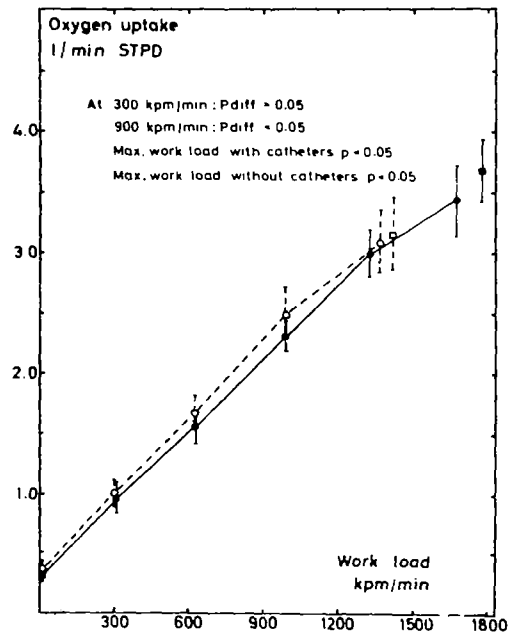
Mean heart rate vs. mean oxygen consumption rate in a group of subjects before (dashed line) and after (solid line) a four-month period of physical training, as discussed in the text (27).



Mean cardiac output vs. mean oxygen consumption rate before (dashed line) and after (solid line) four months' physical training (27).

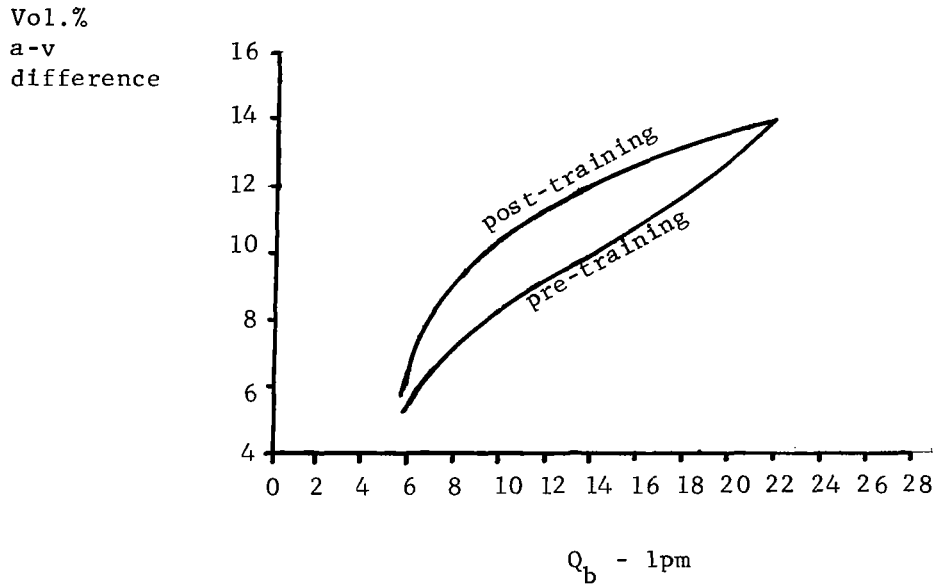


Mean arterio-venous oxygen concentration difference versus mean oxygen consumption rate before (dashed line) and after (solid line) four months' physical training (27).



Mean oxygen consumption vs. external work rate (kilopond-meter/min) before (dashed line) and after (solid line) four months' physical training (27).

The following derived figure illustrates what appears to be one of the most noteworthy changes due to training. The significant a-v difference, as a function of blood flow , indicates that there is an appreciable change in the uptake characteristic in the local vascular beds, presumably in the skeletal muscle bed.



A-V Difference vs. Blood Flow as a Function of Training.

From other data comparing groups of well (and presumably longer) trained athletes with groups of sedentary individuals (for example, see (26)) the resting values of heart rate are lower in trained than in sedentary subjects. When such changes occur in resting performance it is not clear whether resting cardiac output changes as well. In the two groups studied by Bevegard et al., mean resting cardiac output appeared to be somewhat higher in the athletes than in the non-athletes, but the observation is of doubtful significance. About half of the subjects in the Ekblom study showed lower cardiac outputs at rest after training than before. It will require further longitudinal study of individuals to clarify this point.

If the Ekblom curves are used as a reference, one may argue as follows: An athlete's current CV status is achieved by daily sweeping his heart rate over the range of rest, 50-70 bpm, to peaks of 150-180 bpm. However, his average calorie intake, and thus, his oxygen uptake, will only be a moderate amount greater (if at all) than the same weight non-athlete. The non-athlete will sweep his heart rate over the range of rest 70-90 bpm to peaks of 80-120 bpm. If these operating points are located on the Ekblom curves, one can 'guess' that practically no change in average daily cardiac output takes place. Thus, again from the curves, there is essentially no difference in the average a-v difference for the daily average. It remains near 5 vol.% uptake.

What then appears to be the difference in status between the non-athlete and athlete? It seems plausible that the athlete's resting cardiac output is somewhat less than the non-athlete's. However, this simply means that his daily average is higher, because he will likely exercise that day (having likely exercised the previous 20-60 days). Because of that exercise, he has reset the operating point of his skeletal muscle microvasculature. At rest, the athlete may operate with greater peripheral resistance (as if the arteriolar level is operating at a somewhat more constricted state), but at his 'peak' sustained operating range, he likely jams his heart rate up to 180-200 bpm. In fact, one might suggest that the reason for the interminable practice is to be able to achieve the capability for increased uptake at limiting heart rate. Except possibly for events longer than 3 hours, the athlete operates anaerobically during his peak performances.

Thus, at moderate aerobic uptake rates associated with cardiac output, the increased uptake of the athlete for the same blood flow possibly represents also a more constricted state of the capillary beds (i.e., an increase in their 'nutrient' capability). These seem to be plausible conjectures, worthy of further exploration. Thus, it is only to about this vague degree of clarity we have been able to visualize how the operating state of the CV system is adaptively modified. From a systems point of view, on one hand we seem to be confronted with local tissue laws. At rest there is a specific limited range of tissue perfusion (Overlooking the exceptions of glands, skin, and kidney, comparative data in (30), and Goldman suggest a range of 35-130 ml/100 gm./min. for most tissue. The cardiac output they report is somewhat lower, about 18 ml/100 gm./min.); and correspondingly a specific limited range of oxygen uptake. Since the mammalian system operates with a fairly constant high arterial supply pressure, the tissue 'laws' have to represent both the longer term adaptive status of the local vascular bed, and the operational state of the bed. Then, beyond these 'laws' are the status of two local beds, one with a schedule of wide perfusion rates (through exercise), and the other with a limited schedule. Their performance is hardly distinguishable at rest. The 'engineering' systems characteristic that seems to arise, is a moderate difference in uptake at higher perfusion rates. The characterization 'engineering' is appended, because the change in operating status of the two beds seems to reside in some complex 'fine tuning' of the system, which our present knowledge does not permit us to focus on.

A recent, very comprehensive and somewhat illuminating study by Saltin et al. (29) documented the cardiorespiratory adaptations to a 3 week period of bed-rest, followed by a 2 month period of physical training (running and pedalling). During the 3 week bed-rest period the five subjects never were permitted to bear their own weight in the upright posture; maximal oxygen consumption declined 28% from 3.4 to 2.4 lpm, the range in the decline being 20-46%. When the subjects were tested at submaximal loads immediately after the bed-rest period, their load specific oxygen consumptions were unchanged, however. During bed-rest, heart rate during sleep (presumably basal) increased at a mean rate of approximately 0.4 beat/min/day. There was a concomitant decline in heart volume of about 11%, with the largest decline occurring in the two subjects who were previously athletically trained. Resting cardiac output underwent a small decrease during the bed-rest period, but the decrease was not statistically

significant; also, it should be noted that mean resting cardiac output had undergone a further apparent decline by the end of the 53-55 day training period. These small changes are of doubtful significance because of the uncertainty about small differences in the extent to which the subjects were at ease with these measurements throughout the 2 1/2 month study.

All five subjects showed a decrease in maximal cardiac output at the end of the period of bed-rest; the mean decrease being 26%. By the end of the subsequent training period, the three previously sedentary subjects showed a mean 16% increase in maximal cardiac output over the control (pre-bed rest) period; two previously trained subjects showed, respectively a 0 and 7% increase over the control period. In addition to the decreases in maximal cardiac output at the end of the bed-rest period, there was a decrease in (intermediate) work load-specific cardiac output. Comparable decreases with respect to control period were observed during both upright (treadmill) and supine (bicycle ergometer) exercise; thus, more is involved in the decrease than the theoretically adverse effects of the upright posture on venous return.

How are heart and vascular characteristics to be related? Consider these observations just referenced together with the fact that oxygen consumption-specific heart rates were higher than control values after bed-rest, and lower than control values after the training period. It is evident that bed-rest resulted in cardiovascular changes that reduced stroke volume. Thus, as the authors point out, it is not possible to distinguish from their data between a primary myocardial change and a primary vascular change. They chose to call the vascular change in question a defect in the control of venous capacitance; we shall, as discussed later, define the parameter in question as vascular (primarily venous) unstressed volume. The control of venous capacitance (or vascular unstressed volume) is mediated by the sympathetic nervous system, changes in this parameter are manifested by changes in blood volume-specific mean systemic pressure, as defined by Guyton and his colleagues (31). One would expect, were there inadequate control of venous capacitance (vascular unstressed volume), but normal myocardial function, that work load-specific central venous pressures would not differ from the control to the post bed-rest to the post training states. If, on the other hand, this vascular parameter were normally controlled, but myocardial function not, then one would expect to see increases in work load-specific central venous pressure in the post bed-rest state. Unfortunately, but understandably, central venous pressures were not measured, and the heart volume measurements could be made only at rest. It is tempting to guess that the primary defect in the post bed-rest state is an attenuated or ill-coordinated cortically driven entrainment of sympathetically mediated effects on both the myocardium and the unstressed volume of the veins. The presumed interacting effects of these two sympathetic neural effects will be discussed subsequently.

Two other aspects of vascular functioning deserve comment. The total peripheral resistance changes with exercise were not significantly affected by either bed-rest or training. In two subjects, however, immediately post-bed-rest exercise in the upright posture led to a progressive fall in arterial blood pressure and fainting. Presumably the vasodilatation or hyperemia, of exercise developed normally, but with impaired myocardial and/or venous capacitance/unstressed volume adjustments, cardiac output failed to rise in the normal

disproportion to the falling peripheral resistance; the normal disproportion brings about the mild hypertension of exercise. The second noteworthy vascular phenomenon was a small increase in work-load specific arterio-venous oxygen concentration difference. The control A-V oxygen concentration differences at maximum exercise were higher in the two previously athletically trained subjects than in the sedentary subjects, but neither group showed significant changes after bed-rest; the sedentary subjects developed the ability to increase the maximum arterio-venous O_2 concentration difference during the training period from a control mean of 15 vol.% to a post-training mean of 17. (The same magnitude change we showed before.) Since there were only slight work load-specific changes in arterial oxygen content, the increased A-V O_2 concentration difference signifies lower mixed venous oxygen concentrations, which in turn signifies increased oxygen extraction by skeletal muscle and/or an altered distribution of cardiac output in favor of muscle. Serial biopsies of quadriceps muscle throughout the study failed to reveal significant structural changes in vascular architecture. However, since it is a relatively small change in oxygen extraction at issue, the systems or structural concomitants would be most difficult to document.

Noteworthy changes occurred in ventilatory parameters throughout the study. The resultants of ventilation, which were the oxygen consumption specific values of arterial pCO_2 and pO_2, did not differ between control, bed rest or training. Arterial pCO_2 pO_2, and pH dropped slightly as oxygen consumption rose. Although the ventilation-dependent composition of arterial blood was unaffected by bed-rest or training, the character of ventilation changed considerably. After bed-rest, tidal volume increased little with increasing work load, and the increase in respiratory minute volume was brought about almost altogether by increased ventilation rate, which reached the apparent maximum of 50 breaths per minute (bpm) at only slightly over 2 lpm O_2 consumption after bed-rest, whereas 50 bpm was reached at 3 lpm in the control period, and at 4 lpm after training. Diffusing capacity (carbon monoxide) was unaltered by bed-rest or training, although there was evidence for smaller work load-specific alveolar-arterial pO_2 differences, and thus more uniform ventilation perfusion ratios, after training than in either the control or bed-rest states.

It is interesting to contrast the altered ventilatory pattern after bed-rest with what we tentatively presume to be an impaired or dyscoordinated sympathetic stimulation of the myocardium and capacitance vessels. While there is abundant evidence and general agreement that the sympathetic nervous system is activated during exercise, its origins are obscure. It cannot be ascribed simply to the baroreceptor reflexes, since exercise almost always results in moderated elevations in arterial pressure, whereas any simple theory based on baroreceptor reflexes would require a fall in pressure. Nor can it be ascribed to the known chemoreceptors, all of which are in the arterial stream, since arterial pCO_2 need not (and in the Saltin study, did not) rise with exercise, and the slight falls in arterial pO_2 were too small to be an effective chemoreceptor stimulus. For the same reasons, arterial chemoreceptor drive on ventilation is singularly unconvincing. Yet large changes occur in both ventilation and sympathetic stimulation of the cardiovascular system in exercise.

It is as tempting in 1970 as it has been for the past half-century or more

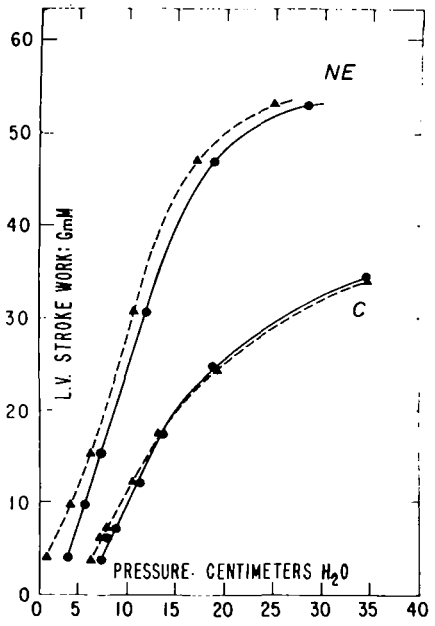
to speculate on a chemoreceptor which monitors the mixed venous stream, but no such receptor has ever been found which is capable of activating such large changes in cardiovascular - respiratory function (see, for example (32).) In the respiration literature there is the suggestion of Dejours (33) about "fore-brain drive" on ventilation, which we may translate and expand to call a cardio-pulmonary control algorithm, which we would postulate to be an inescapable neurophysiological concomitant of volitional movement. There is, on the vascular side at least, some skimpy evidence for this already in Orville Smith's observation in monkeys that electrical stimulation in the motor cortex (area 4), which was capable of eliciting discrete coordinated arm movements, also elicited decreases in vascular resistance, in the stimulated limb (? via sympathetic cholinergic vasodilators) but not in the contralateral limb, which persisted after the movements (and their local metabolic concomitants) were blocked by curare (34).

Another illustrative example is the work of Miller, who was able to select an autonomically mediated vasomotor response - vasodilatation in one of a test rabbit's two ears - by appropriate conditioning(35). The implication of Miller's achievement is two-fold. First, like Smith's work, it shows that autonomic vasomotor influences can be sharply localized and need not be diffuse. Secondly, it shows that autonomically mediated effects can be "learned". It is to be expected that a considerable increase will take place in reports of such effects. (see for example (10)). The assumption seems to us plausible and certainly deserving of close experimental scrutiny that the control structure for volitional movement includes algorithms for the activation of both ventilatory control mechanisms and the cardiovascular sympathetic nerves, which govern not only the patterning of heart rate (in conjunction with the vagus) but also changes in resistance and unstressed volumes in the vascular tree, and the pumping characteristics of the heart. A corollary of this general notion is that training and de-training effects are partly reflected in the coordination and dyscoordination, respectively, of these vegetative functions; in other words, the complex factory algorithms admit to training and de-training.

The editorial prelude by Asmussen (36) to a recent symposium on the physiology of muscular exercise speaks of the "work factor" - that undefined but large stimulus to the cardiovascular, ventilatory, and thermo-regulatory systems in exercise. That choice of terms is strongly suggestive of a chemical, thermal, or mechanical factor - perhaps too strongly so. We would suggest that, just as the CNS has an algorithm for voluntary muscle co-ordination in various patterns of movement, so it has associated algorithms for coordinating the "vegetative" functions which support volitional movements. Is this a useful more or less new idea, or is it only a verbal smokescreen which will obscure rather than clarify? It depends upon whether this notion provokes new experimental approaches; unfortunately, since it is a problem in higher nervous activity and organization it may be too subtle an idea to test at present by direct operational means. (For example, we have already implied that both exercise, emotional state, learned behavior, and behavioral complexes changes the microvascular state. It is not within the state of the art to pursue this question experimentally.)

6. An Engineering Definition of Systems Relations Governing Cardiac Output.

Systems relations, which can arise from either theory or phenomenology, will inevitably include a number of empirical parameters, and these have to be accounted for. That accounting is or will be in terms of mechanisms operating at a hierarchically lower level than the level of function described by the systems relations. For example, a parameter commonly used in the physiological literature for the intact heart is "contractility"; unfortunately, it has many meanings, but one in common use is to denote (connote is perhaps the more accurate term) the position of a curve in the end-diastolic ventricular volume - external ventricular work plane. An increasing contractility is shown in the following figure. One of the commonly accepted tasks of cardiac physiology is to come to an understanding of the underlying myocardial mechanisms which account for the transition between the curves in the figure.

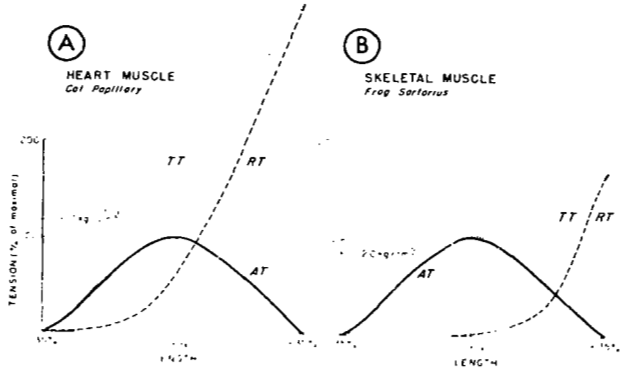


Left ventricular stroke work (for definition, see text) versus left ventricular end-diastolic pressure (solid lines) or versus left atrial pressure (dashed lines). The pair of lines labelled "C" was taken in the control state, and the pair labelled "NE" was taken during an infusion of norepinephrine (39).

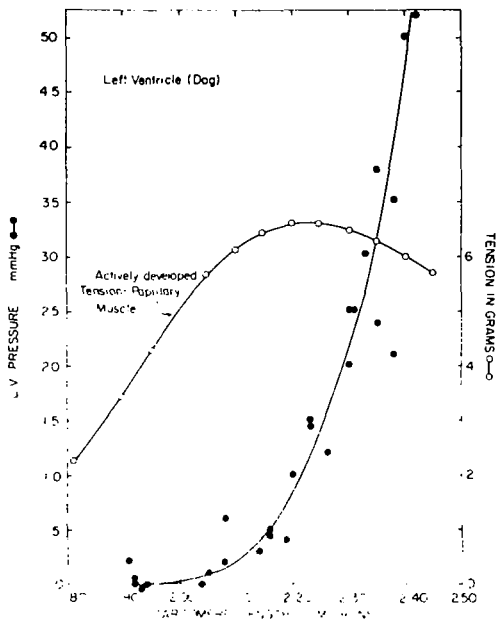
In general terms, it will be our epistemology to choose a hierarchical level of physiological function, to note kinematic relations obtaining at that level, to formulate systems relations and their concomitant parameters, and then utilize the parameters as the transitions to the underlying mechanisms operating at a lower hierarchical level. The art in this science lies in the choice of paramaters. Where they are vague and without an explicit metric, redundant, or ultimately not subject to independent derivation, 'proof', the system equations in which they arise are poorly chosen.

(a) An engineering definition of cardiac function. Let us turn to consider the heart, as both a specific example of the above points and a beginning of our discussion of cardiovascular system. The history of cardiac physiology is closely related to the development of skeletal muscle physiology, for there has been a continuing attempt since the turn of the century to describe cardiac function in terms which are (or were) in use in the description of the mechanics and energetics of skeletal muscle.

Otto Frank was one of the pioneers in this effort. He succeeded in the 1890's in showing that the frog's myocardium displayed a relation between fibre length and isometric tension which qualitatively resembled that of skeletal muscle, which had been shown earlier in the 19th Century. Representative modern data (37) using papillary muscle instead of the whole ventricle, are as shown in the next two figures. A similarly shaped curve exists for skeletal muscle, although the peak developed tension is approximately twice that in papillary muscle, per unit area.



Effect of length on resting tension (RT) and tension developed during contraction (TT) in cat papillary (myocardial) muscle (A) and frog sartorius (skeletal) muscle (B). The heavier lines, labelled AT, plot the difference between TT and RT, and represent the amount of tension (active tension) developed upon stimulation of the muscle (37).



Plot of active tension (open circles; right hand ordinate scale) versus sarcomere length in the cat papillary muscle. The solid circles show the relation between left ventricular end-diastolic pressure and sarcomere length in the dog heart. Note that the peak in active tension occurs at the sarcomere length which approximately coincides with the upper limit of physiological left ventricular end-diastolic pressure (37).

Just prior to World War I, Starling and his pupils re-investigated the phenomena Frank had elucidated, but made several modifications which gave their findings an apparent applicability to many clinical problems. They studied the mammalian heart under simulated working conditions, rather than limiting their attention largely to isovolumic function as Frank had done. They showed that the energy produced in contraction (which could be judged by the product of cardiac output and developed pressure) was in proportion to the volume of the heart at the time each contraction (systole) began. This volume is commonly called end-diastolic volume (EDV). Since EDV changes in the same direction as myocardial fibre length, Starling proposed what came to be known as Starling's Law of the Heart: "The law of the heart is therefore the same as that of skeletal muscle, namely that the mechanical energy set free on passage from the resting to the contracted state depends on the area of 'chemically active surfaces', i.e. on the length of the muscle fibre. This simple formula serves to 'explain' the whole behaviour of the isolated mammalian heart.".(38)

Starling's work was formulated into a view of the regulation of cardiac output (Q_b), which persisted until mid-century. That view regarded the heart as if it were an energized turnstile, which accepted incoming venous blood at the prevailing Q_b and raised it to the prevailing regulated arterial pressure. As the work involved was proportional to Q_b, the heart automatically increased or decreased its EDV in keeping with the energy required. This view placed the emphasis on vascular phenomena as the determinants of Q_b.

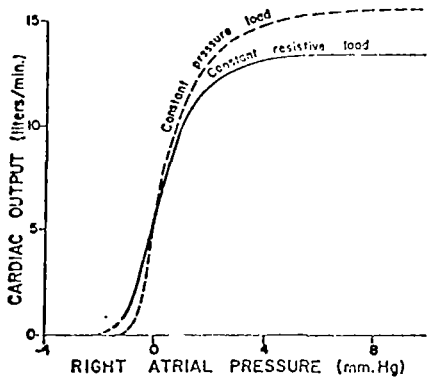
In the mid-fifties, two concurrent lines of investigation made it apparent that the Law of the Heart would have to be modified. One line was a group of studies on cardiac dimensions in intact animals and man in the rest-exercise transition; these studies in general failed to reveal the expected dimensional changes. More recent work showing the cardiac dimensional changes in exercise was noted in the previous section. The other line was studies on neural and humoral influences on the isolated heart, which were pioneered by Sarnoff and his pupils (cf (39)).

Sarnoff's studies differed from Starling's in that, whereas the latter had studied the heart (and lungs) altogether isolated from systemic influences and had had independent control over Q_b, Sarnoff contrived to monitor cardiac function in intact (though anesthetized and extensively operated upon) dogs. In Starling's experiments, the heart and lungs remained in situ, but blood was pumped from the heart through an extra-corporeal shunt into a reservoir whose height above the right atrium provided the experimenter with independent control over Q_b; the rest of the dog received no arterial blood and was thus dead and without influence on the heart. Sarnoff allowed the monitored heart to perfuse the body, and varied Q_b by hemorrhage or transfusion, and also either reflex or directly controlled (or attempted to control) autonomic or hormonal influences on the heart. His studies showed that there was not a unique relation between EDV and the external work of the heart. Sarnoff's work, which is summarized in his contribution to the Handbook of Physiology, and the work of others may be summarized by stating that the description of cardiac function requires at least a 7 - space, whose axes are:

1. External work
2. EDV
3. Cardiac sympathetic (accelerator) nerve firing rate
4. Heart rate (f)
5. Catecholamine concentration in coronary arterial blood
6. Vagal (parasympathetic) nerve firing rate
7. Aortic (ejection) pressure (also known as afterload)

The above list is ordered according to our guess as to the relative quantitative importance of factors 3-7 in influencing the two-dimensional "Law" of Starling.

Guyton, in formulating a scheme of Q_b regulation (31) which will be discussed in some detail below, has extrapolated a different functional description of the heart from this literature. He chooses a Q_b - right atrial pressure (RAP) relation as the "basic" two-dimensional law. His relation is shown in the following figure.

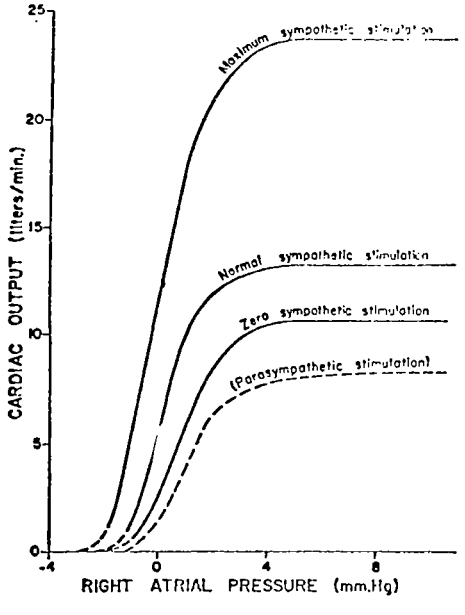


Cardiac function curves plotted on the assumption of constant resistive load (solid line) or constant pressure load (dashed line). The discussion here treats the constant resistive load case (31).

The RAP is a transmural pressure across the wall of the right atrium. The justification for shifting from EDV - a ventricular value - to a transmural atrial pressure lies in Sarnoff's and others' subsequent (cf.(40)) demonstrations that the diastolic pressure - volume relation is an invariant and that the atrium and ventricle constitute a common chamber during diastole. The constancy of the atrial pressure-ventricular end-diastolic volume relation presumes a normal atrio-ventricular valve and a normal pericardium, so that these questions must be re-opened when there are pathologies involving those structures.

The choice of the right atrial pressure as the variable used to describe the intact heart stems from Sarnoff's work showing clearly that, in the normal heart, the right is the limiting ventricle in respect to the performance of external work. Under a given set of systemic influences, the approximate peak work of the left ventricle is 8-10 times that of the right. The mean afterloads on the two ventricles differ (L/R = approx. 5) and it is Guyton's estimate that the peak flows realizable by each ventricle are approximately equal. The issue is further influenced, however, by the fact that left atrial pressure (LAP) may make a significant contribution to the afterload on the right ventricle, so that as Q_b is increased toward the maximum, the $Q_{b_{max}}$ for the right may reasonably be expected to drop below the $Q_{b_{max}}$ for the left. The point is an important one in in that $Q_{b_R} > Q_{b_L}$ is a rapidly lethal condition.

Guyton's description (31) treats the variables 3-7 in the above list as parameters on the RAP-Q_b relation. An example is the influence of the cardiac sympathetic nerves, shown in the following figure.



Effects of changes in sympathetic or parasympathetic stimulation on the cardiac function curve (31).

A useful simplification of these multidimensional relations is possible. Note that the curves have similar shapes, and they approximately obey the following empirical rules

1. $Q_b = 0$ when RAP = 0

2. Q_b = max at RAP = 8

3. Q_b = 0.4 max at RAP = 4

4. The curve is concave upwards in the range $0 < RAP < 4$, and concave downwards in the range $4 < RAP < 8$.

It is therefore convenient to coin a descriptive parameter, which we can call conditional peak flow (CPF), and write down the following empirical relations which D. A. Gall has derived:

$$Q_{b_L} = CPF_L \; y^2 \; (1.4 + 1.2\,y - 1.6\,y^2)$$

$$Q_{b_R} = CPF_R \; x^2 \; (1.4 + 1.2x - 1.6x^2)$$

$$Q_{b_R} = Q_{b_L} \text{ for the intact heart}$$

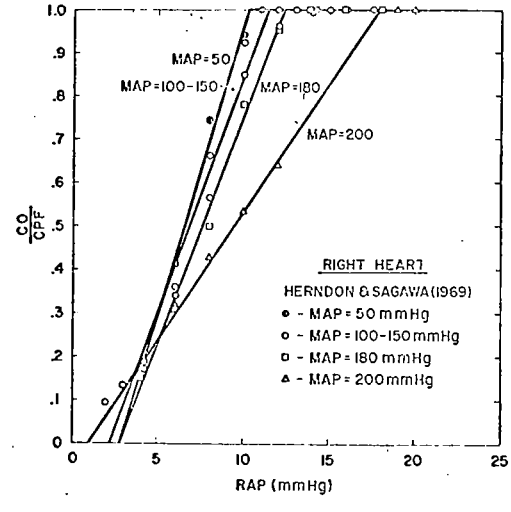
$$\text{where } y = \frac{LAP - IPP - 2}{16}$$

$$x = \frac{RAP - IPP}{8}$$

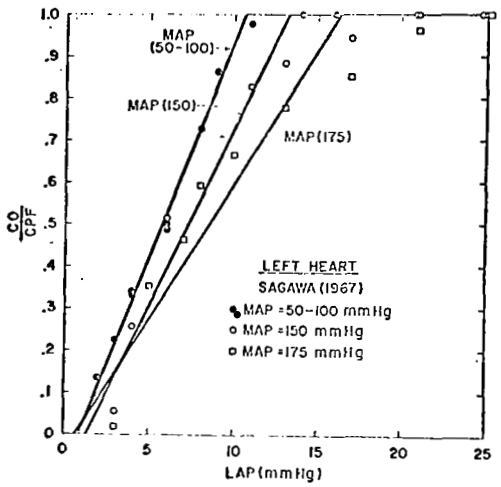
and IPP = intrapleural pressure, which is normally about 4 mmHg below atmospheric in the closed-chest dog or man. The differences in y and x describe the different diastolic capacitances of the two ventricles.

This manner of expressing the function of the heart has both advantages and difficulties. The chief advantage is to have reduced the pumping action of the intact heart to a single empirical parameter, CPF, which has explicit dimensions (flow). The influence of variables 3-7 on CPF can, in turn, be described in a series of two dimensional relations. The second advantage is that, in theory, a single measurement of Q_b at RAP > IPP + 8 would provide a description of the whole curve. The assertion has only retrospective experimental validation, and points up the chief difficulty with this view as of this writing. Guyton's curves after all come ultimately from the large mass of data collected by Sarnoff in a technically difficult experimental preparation in which control was limited. The data were smoothed and filtered for the purposes of Guyton's modelling. Then we have further filtered by imposing a single shape. Whether the resulting view, based on the single parameter, CPF, has heuristic value remains to be seen.

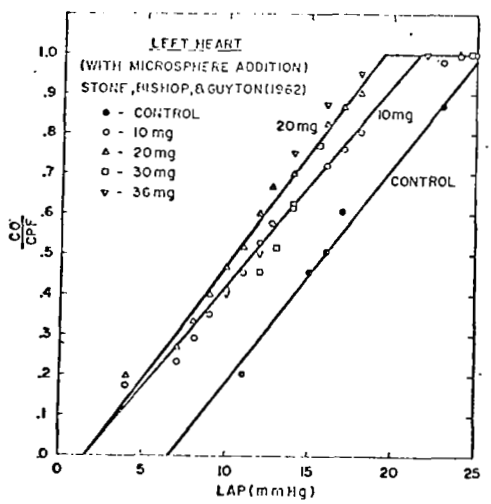
As a preliminary test, the next three figures show data from two relatively recent sources, which were sufficiently complete to be examined against the notion of the single parameter model. Sagawa's (41) and Herndon and Sagawa's (42) data on left and right heart function conform reasonably well to the single parameter model as long as systemic arterial pressures were held within the physiological range. An earlier paper by Stone et al. (43) shows a much less steep slope than Sagawa found for the left heart; also there is an unexplained leftward shift in the curve after embolization of the coronary vascular bed by injections of 10-36 mg of microspheres. However, between 10 and 36 mg of microspheres, the conditional peak flow declined from 3.5 to 2 lpm, indicating that the single parameter serves to describe a range of sub-normal cardiac function. We have been struck by the fact that, despite much animal experimentation, there is such a very limited amount of useful data available. A much more extensive amount of data is available that serves to indicate qualitatively how factors 3-7 on the above list influence cardiac function, but the quantification of these has been sparse.



Cardiac output (CO), in relation to conditional peak flow (CPF) versus right atrial pressure (RAP) at different mean arterial blood pressures (MAP). Data taken from (42).



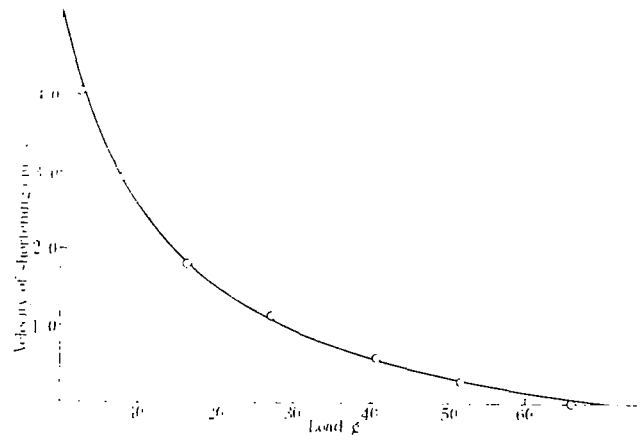
Cardiac output (CO), in relation to conditional peak flow (CPF) versus left atrial pressure at different mean arterial blood pressures (MAP), (41).



Cardiac output (CO), in relation to conditional peak flow (CPF) versus left atrial pressure (LAP) before (control) and after embolization of the coronary arterial tree with 10-36 mg of microspheres (43).

The point of gaining a parameter descriptive of cardiac function from a single measurement or set of measurements is important, for the view that we have been left by Sarnoff is almost impossibly cumbersome when one seeks to quantify cardiac function in one or another physiological state. The ventricular function curve (VFC) occupies a unique position in the end diastolic ventricular pressure - external work per stroke plane, and is constructed out of many pairs of measurements which of necessity have to be collected over a period of many beats under difficult to maintain conditions of stationarity. Sarnoff used a parameter, contractility, to give qualitative indication of the position of the VFC in the plane. "When, from any given end diastolic pressure or fiber length, the ventricle produces more external stroke work and more external stroke power (stroke work per systolic second) an increase in ventricular contractility is said to have taken place, and vice versa." ((39), p. 490). This is a classically bad coining of a physiological parameter, for, at the very least, it lacks explicit dimensions. The usage of the term contractility is further complicated by multiple definitions, applied to skeletal muscle, and to the smooth muscle of various viscera, including blood vessels.

Within the past half-dozen years, a quite different approach to describing the heart has been launched, which has commanded wide attention. It represents an attempt to strike an analogy between cardiac and skeletal muscle at the level of a relation which has come to be accepted, together with the length tension relation already noted, as a basic mechanical property of muscle. This is the force-velocity relation, which for skeletal muscle is as shown in the following figure. Such data are obtained from measurements on isotonic



Force-velocity relation for frog sartorius muscle, stimulated tetanically at $0^{o}C$ (45).

shortening velocity under a wide range of loads. It is a characteristic which can be shown for single skeletal muscle fibres, large skeletal muscle masses, and papillary muscle, which is a readily accessible piece of myocardium that has parallel fibres. A statement equivalent to the force-velocity relation is that, over a wide range of loads, power development is approximately constant. Since the duration of a single contraction is approximately constant, or, more commonly, when the relation is worked out during tetanic stimulation, this relation signifies that muscle releases more useful energy as its load increases. In fact, it releases more total energy as load increases, as Fenn's heat measurements first showed in 1923 (44). The Fenn effect is one of the remarkable properties of muscle.

One of the relatively short-lived triumphs of skeletal muscle physiology was the following line of reasoning, which implanted the force-velocity relation as a fundamental property of muscle. A. V. Hill (45) noted from the heat evolution by skeletal muscle that the "extra" energy associated with increasing loads was a linear function of tension development:

$$\text{Rate of "extra energy"} = b\ (P_o - P)$$

where P_o is isometric tension (an initial length-specific muscle characteristic)

P is developed tension
b is a coefficient

Hill noted, from heat measurements, that a heat, ax, was evolved in association with shortening through distance x, and that the

$$\text{Rate of "extra energy"} = av + Pv$$

where v = velocity

Therefore,

$$(P + a)\ (v + b) = (P_o + a)\ b \quad ,$$

the equation for a rectangular hyperbola, whose asymptotes are

$$P = -a \text{ and } v = -b$$

Thus, it appeared possible to obtain 'a', a thermal constant, from mechanical measurements. That apparent ability to relate thermal and mechanical events gave a special significance to the force velocity relation in muscle physiology. Also, for the skeletal muscle, the coefficients of the equation apply to muscle from a wide range of species. In recent years, however, it has become evident that the heat measurements were in error, to the extent that the identity of the thermal and mechanical properties has been lost (46). Nevertheless, the attempt has been made to try to compute myocardial force-velocity relations from data on single beats of the intact heart, and, together with a

geometrical model of the heart, to use these as an alternative to the VFC as a description of "the contractile state", or the pumping capability of the heart.

The assumptions and measurements required to arrive at a force-velocity relation for the heart are detailed in a recent monograph by Braunwald, Ross and Sonnenblick (40). Such curves can be computed and they shift about in keeping with a priori qualitative expectation in response to various inotropic interventions. Factors which, in Sarnoff's definition would increase contractility, effect an increase in both v_{max} and P_o, whereas changes in fibre length influence only P_o.

The experimental procedure in determining these curves and the interpretation applied to them are based on a idealized model of muscle which presumes a mass-less contractile element acting in series with a mass-less elastic element. The contractile element is assumed to be freely extensible at rest, and inextensible during stimulation. A second elastic element is presumed to be parallel with the contractile element (CE) and the series elastic element (SE). The parallel elastic element (PE) accounts for the extensibility characteristics of the muscle at rest. This model was developed for skeletal muscle by Hill in 1938 (45), and is presumed applicable to myocardium. In the calculation of the force-velocity curve, it is the velocity of shortening of the contractile element which is sought. Thus, the experiments must be done under circumstances in which the tension in the SE balances the load. P. so that the SE is neither lengthening nor shortening, and that the observed velocity of contraction is the velocity of shortening of the CE. In Hill's terminology, stimulation of the muscle produces a characteristic time course of tension development by the CE. That tension is called the active state, which has some characteristic time course. In this view, one can observe the time course of the active state by forcing the SE into rapid equilibrium with the load. One means of forcing this rapid equilibrium has been to quickly stretch the muscle just after its stimulation, adjusting the stretch to bring the tension in the muscle as near as possible to the value the muscle would develop spontaneously by a shortening of the CE and a lengthening of the SE. If the SE can be forced to equilibration with the CE, then the time course of the active state is revealed either by the time course of isometric tension (velocity equalling zero), by velocity-tension pairs, or by maximal velocity at zero load. Thus the position of the CE force-CE velocity curve indicates active state intensity. There will be a surface in force-velocity-time space which reveals the time course of active state, which might be rephrased to be called the maximal contractile potential of the fibre.

The difficulty in applying this line of reasoning to myocardium is that it is not possible to tetanize cardiac muscle. In skeletal muscle, one can repetitively stimulate (tetanize) and presumably bring active state up to some sustained intensity and then have time to make the CE force-CE velocity measurements without the problem of time variations in active state intensity. During a single twitch (say 200 msec in myocardium) the active state changes rapidly in time and thus the problem of equilibrating the SE and making the velocity measurement becomes troublesome. The probability is high that force-velocity curves computed from data on multiple contractions represent multiple

active state intensities slurred over.

It may be objected, too, that the whole conceptual approach of the active state of an hypothetical "pure" CE and the simple lumped-parameter mechanical model has been too far superseded by knowledge of muscle ultrastructure and more recent theories of muscle contraction mechanisms. We hazard the comment that it is odd to see myocardium coming to be described in terms of Hill's 1938 model just as that model was being supplanted by the much more nearly isomorphic views which are derivative from the Huxley sliding filament model (cf. (47)).

Finally, there is the problem for those working at the organ level that the hypothetical force-velocity curves are too remote from the functioning of the intact heart to be of much use in describing the heart as a pump.

For those reasons, we have provisionally elected to use conditional peak flow as a pump parameter, recognizing that, while it is not yet obviously accounted for in terms of the underlying myocardial muscle mechanics, it has some promise of stimulating quantitative experimental testing of the heart, and of characterizing the heart in useful terms for a vascular circuit analysis, to which we now turn.

(b) The vascular circuit - an engineering view.

The vascular circuit presents an unusual hydraulic conduit. There is a low capacitance arterial tree terminating in high resistance vessels known as arterioles (15-35μ diameter). Beyond the arterioles is the capillary network, the one portion of the vascular circuit that permits radial as well as axial flow. The 'typical' capillary is about 5 μ in diameter, and perhaps 100-150 μ long. Beyond the capillaries are low resistance high capacitance venules and veins. The overall dimensions of the vascular tree are such that its unstressed volume is approximately 85% of the blood volume. The overall capacitance of the vascular tree, at volumes greater than the unstressed volume, is said to be in the vicinity of 100 ml per mmHg for men (cf. (31)).

Which of these parameters are system variables? In the short time scale (e.g., 400 sec.), blood volume and all the capacitances appear to be constants, with arteriolar, venular and venous resistances variable. Also, an important variable appears to be the unstressed volume of the circuit, the major fraction of which resides on the venous side.

One can write compliance equations for the major segments of the circuit, of the form:

$$V_i = C_i P_i + (V_o)_i$$

where the subscript 'i' refers to the segment in question, C is its capacitance (i.e., its compliance), P is its transmural pressure and V_o is its unstressed volume, and corresponding resistance equations,

$$P_{i-1} = R_i Q_i + P_i$$

where the subscript i-1 refers to the next upstream (antegrade to the direction of flow) segment, and Q_i is the flow through the segment.

A plausible set of parameter values for major segments in man, in accordance with Guyton's modelling, might be:

Segment	C - 1/mmHg	R - mmHg/lpm	V_o - liters
systemic arteries	.002		.5
		16	
systemic capillaries	.005		.15
		4	
systemic veins	.06		2.0
		0.2	
pulmonary arteries	.001		.4
		3.2	
pulmonary veins	.012		.65
right atrium and great veins of thorax	.02		.48

Blood volume = 5 liters

Overall capacitance = 100 ml/mmHg

This simplified model of the physics of flow through the vascular circuit has the properties of the vascular system that Guyton and his colleagues have emphasized in their scheme of short-term cardiac output regulation. This scheme is extensively developed in Guyton's monograph, "Cardiac Output and its Regulation" (31).

Disregarding the question of specific numerical values, the present view differs from the one put forth by Guyton in several respects. First, cardiac function is described by us with a single parameter for each ventricle, conditional peak flow, whereas Guyton has no explicit parameter to describe the influences of such factors as the autonomic nerves, blood catecholamines, rate, and afterload on the pumping ability of the heart. Second, we treat the

unstressed volume of the venous tree as one of the important variable parameters by which the cardiovascular system is forced, whereas Guyton chooses a virtual pressure, the mean systemic pressure, to be the variable parameter representing the same influences on the vascular tree.[1] The mean systemic pressure is the equilibrium pressure at zero flow in the vascular tree. In general, pressures make poor choices as system parameters; in this instance, mean systemic pressure changes can arise from blood volume changes, vascular capacitance changes, or unstressed volume changes. The latter two arise from changes in autonomic (sympathetic) nervous activity or from changes in compressive effects of skeletal muscle in which portions of the vascular tree are imbedded. In contrast to these rapidly (10-100 sec) changing influences, physiological blood volume changes occur slowly over a period of days, although, pathologically or experimentally, blood volume can be changed with arbitrary rapidity. The difficulty with venous capacitance and unstressed volume as system parameters is that they are effectively unmeasurable directly and the better views we have of them have come from Guyton's measurements of mean systemic pressure at constant blood volume. However, unstressed volume changes can be inferred from controlled studies on perfused vascular beds, an excellent example of which is the study of Mellander (48).

The resulting physical-physiological model has superimposed upon it two major regulatory mechanisms, the baroreceptor reflexes and the metabolic autoregulation of vascular resistance. Each deserves comment.

The baroreceptor reflexes influence arteriolar resistance (called systemic arterial resistance in the table) in such a way that arterial pressure is held relatively constant. There are four major groupings of stretch receptors in the arterial tree:

(1) Right carotid sinus
(2) Left carotid sinus
(3) Aortic arch
(4) Origin of right subclavian artery

1. The two parameters are related by

$$(V_b - V_o)\left(\frac{dp}{dV}\right)_v = \Delta p_{ms}$$

V_b = blood volume

V_o = unstressted volume

$\frac{dp}{dV}$ = compliance of the total vascular tree

Δp_{ms} = mean systemic pressure

We might incline to regard V_o and $\left(\frac{dp}{dV}\right)_v$ as species and weight specific. In that case the regulation of blood volume V_b, or excess blood volume $V_b - V_o$ and of mean systemic pressure become equivalent physiological questions.

Afferent nerves from each receptor site synapse extensively in the brainstem and influence : (a) sympathetic nerve impulse traffic to both heart and systemic blood vessels and (b) vagal nerve impulse traffic to the heart. The cardiac influences are such that, as pressure at the receptor site falls, heart rate increases and conditional peak flow increases (over and above the effects attributable to the rate changes (49)). The blood vessel changes are such that, as pressure at the receptor site falls, arteriolar resistance rises. Other blood vessel influences - capacitance or unstressed volume changes - have been hard to document. There is evidence (50,51) for baroreceptor reflex influences on venous capacitance or unstressed volume but it is clear, as well, that the baroreceptor reflexes have little or no ability to raise cardiac output above the normal metabolism-specific level (cf.(23)). Therefore, it would appear that, if there are appreciable system-wide influences of the baroreceptor reflexes on venous capacitance or unstressed volume, it must be offset by concomitant venous resistance changes so that cardiac output is little influenced. At the same time, however, it does appear that baroreceptor reflexes act to minimize the reduction in cardiac output with hemorrhage. Thus, when cardiac output falls below a metabolism - specific level, the qualitative character of the baroreceptor reflexes seems to change so as to include significant flow effects (23).

The arterial pressure regulation afforded by the baroreceptor reflexes has been the subject of a number of experimental analyses which follow linear or elementary nonlinear control theory (24, 52-54), although one of the first studies of the reflex by Koch in 1931 (55), after its discovery by Hering in 1927, involved a seemingly unwitting quantitative analysis of the open loop carotid sinus reflex. Koch isolated one carotid sinus from the arterial tree and subjected it to controlled hydrostatic pressures while noting the concomitant systemic arterial pressures. The other sinus was acutely denervated and both vagi were cut, so as to deafferent the two other baroreceptor sites. The results established an open loop gain of the reflex change of systematic pressure per unit change of carotid sinus pressure of 1-3 in the region of physiological blood pressures. However, the significance of this finding went uninterpreted until Scher and Young's essentially repetitive studies in 1963 (52). It is seemingly only an historical coincidence that Koch (1931) carried out his extensive studies on the open loop reflex approximately coincident with the publication of Minorsky's and Nyquist's pioneering works on the stability of technological control systems based on open loop performance.[1] Even with the analysis borrowed from linear control theory, however, the interpretation of experiments on the open loop system is not straightforward because the interactions between the four baroreceptive sites are not clear. Sagawa and Watanabe (56) and Stegemann (53) have examined the interactions between left and right carotid sinus reflex actions and found them approximately additive (Stegemann drew a different conclusion from his data, but given its scatter one is hard put to argue that there is other than an approximately additive interaction). No one has succeeded in gaining experimental control over the

[1] Others might regard this as additional reinforcement of the concept of the technological 'agenda of history'.

pressures at all four sites, because the aortic arch and, to a lesser extent, the origin of the right subclavian are major conduits. Control of the pressure requires extensive interference with cardiovascular function. A recent example of work on the aortic arch baroreceptor is the work of Allison et al. (57) who found an open loop gain in the change of mean arterial pressure per unit change in aortic arch pressure of 0.5 or less. Thus, not only the interactions among these parts of the baroreceptor reflex system, but also the strength of their combined contribution to blood pressure regulation is unmeasured.

Four nonlinearities in the carotid sinus reflex deserve mention. The first is saturational and occurs on either side of an operating range of about 60 to 160 mmHg. The second is that the open loop gain is cardiac output-dependent; the reflex affects principally arteriolar resistance in a manner which is, to first approximation, independent of cardiac output. Thus, the strength of the influence of carotid sinus pressure on systemic pressure is proportional to cardiac output. The third is that the reflex is dynamically asymmetrical (24), in that systemic pressure falls more rapidly following a sinus pressure increase than it rises following a sinus pressure decrease. The fourth, and perhaps most interesting, is that the gain of the reflex is inversely related to the high frequency content of carotid sinus pressure. The data of Ead, Green and Neil (58) show this property - although those authors did not perceive it - with normally pulsatile carotid sinus pressure. The change in mean systemic pressure per unit change in carotid sinus pressure was larger than with nonpulsatile carotid sinus pressure. Levison et al. (24) showed that, during low frequency sinusoidal variations in carotid sinus pressure, if a high frequency low amplitude pressure variation were superimposed, systemic pressure not only sought a lower mean level, but also showed approximately half the low frequency gain that prevailed without the superimposed high frequency pressure variation. An interesting implication of this fourth nonlinearity is that the decline in pulse pressure that accompanies hemorrhage may act, to some extent at least, to increase the reflex gain and offset the drop in gain with falling cardiac output.

There is an immense literature on the baroreceptor reflexes, most of which concerns short term (1-3 hours) experiments. An interesting long-term change in baroreceptors is their apparent adaptation to sustained hypertension. After a relatively few days of hypertension, the sensitivity range of the baroreceptor shifts, from being approximately centered in the physiological pressure range, to being approximately centered in the new pathological pressure range (59). The mechanism of this shift in the operating range of the baroreceptors is unknown. Its implication, however, is that long-term pressure regulation must be found among other mechanisms.

The metabolic 'autoregulation' of vascular resistance is a prominent characteristic of the following vascular beds: skeletal muscle, myocardium, and brain. In the brain, vascular resistance is inversely related to the pCO_2 of cerebral blood (60). Thus if cerebral flow falls below some cerebral metabolism - specific level, the prevailing pCO_2 in the fluids surrounding cerebral arterioles will rise, with resulting vasodilatation resistance fall, and

compensatory flow increase. In skeletal muscle and myocardium, the effects of CO_2 are not sufficiently strong to account for the tendency of these two vascular beds to vary their resistance as their metabolic rate - blood flow relation changes. The classical terminology, based on studies of vasomotion in skeletal muscle, is that the fall in resistance (and subsequent increase in flow) which accompanies arterial occlusion is called a reactive hyperemia; the fall in resistance which accompanies an increase in the tissue's metabolic rate gives rise to the so-called active hyperemia. These responses were discovered in the 1880's by Gaskell, working in Carl Ludwig's laboratory, but to date no adequate mechanism for either response has been documented. These responses appear to be altogether independent of the autonomic nervous system and so are called "local", "metabolic", or "auto" regulatory. There are no known pharmacological blockers of the response. Whatever the mechanism, there is some intimate relation between vascular resistance and metabolic activity of muscle such that as metabolic activity increases, vascular resistance falls. It is a vasomotor response of fundamental significance, which insures that the local oxygen cost of centrally driven exercise be met. In the myocardium a seemingly analogous mechanism insures that coronary blood flow convects an oxygen supply adequate to meet the energy cost of maintaining the cardiac output, while at the same time holding tissue pO_2's at acceptable mean levels. An important point about the regulation of vasomotion in skeletal muscle is that, where sympathetic vasoconstriction and metabolically induced vasodilatation may conflict, the metabolic vasodilatation has priority (61,62). Thus it appears to be true that in exercise, the simultaneous sympathetic stimulation (a concomitant of the central 'command algorithm' in exercise) and metabolically induced vasodilatations interact to reduce overall vascular resistance in skeletal muscle, while at the same time reducing unstressed volume. These two effects facilitate venous return, and thus are of major importance in elevating cardiac output from its resting value to values 3-5 fold greater in exercise. The metabolic vasodilatation may be thought of as providing a follower characteristic in cardiac output regulation, for it lags in time behind the onset of increased metabolism. When the central 'command algorithm' coordinates properly, the sympathetically mediated effects in vascular unstressed volume and conditional peak flow can serve to provide anticipatory increases in cardiac output, which together with a possible sympathetically mediated vasodilatation in soon-to-be-active muscles and constriction in gut and kidney can serve to provide anticipatory redistribution of the increased cardiac output. The fortuitous observation of Saltin et al. (29) that several of their subjects became hypotensive and fainted during post bed-rest exercise suggests that the follower characteristic outran the normally anticipatory adjustments which depend on the central 'command algorithm'.

The chemical (or thermal) coupling between muscle activity and vascular resistance has been a subject of much study for many years. Recently, Skinner and his associates (63, 64) have investigated the effects of selective changes in oxygen tension, potassium concentration, and osmolality in arterial blood perfusing the canine gracilis muscle. They showed that, while reductions in arterial pO_2 would result in a resistance fall, it was possible to elicit only a minor fraction of the maximal response observed when the gracilis was caused to do work. Similar results were observed with increases in

arterial potassium concentration that were chosen to mimic the small increases in potassium concentration observed in venous blood draining the working gracilis muscle. However, when there was both a reduction in arterial pO_2 and a rise in arterial potassium concentration, near maximal reductions in resistance were observed (63). More recently (64), Skinner has added a third factor, a small increase in blood osmolality, which also occurs in venous blood from working muscle, and has shown that it too interacts with hypoxia and hyperkalemia to produce changes in resistance of comparable magnitude and time course to those observed after the onset of muscle work. A fourth factor, which ought to be examined critically and quantitatively, is the vasodilatory effect of temperature. It may well turn out that the hyperemia of exercise, which is such a large and important feature of short-term cardiovascular regulation, is the resultant of a number of interacting but individually small changes in the composition of interstitial fluid and plasma locally in exercising skeletal muscle. The problem of quantitating the role of these various candidate-hyperemic factors is complicated by the presence in muscle of a variable number of direct a-v shunts, or thoroughfare channels, which in effect contaminate the venous blood with arterial blood and so lead to an underestimate of the compositional changes locally in the region of the vasomotor (largely arteriolar) changes. Skinner et al. do not comment on this problem, but it is fundamental to the attempt to quantitate the role of various locally produces hyperemic factors. It is axiomatic that only vessels small enought not to have vasa vasora are subject to active or metabolic vasodilatation. Larger vessels, which possess vasa vasora, have arterial blood convecting through their walls and are effectively isolated from the chemical consequences of activity changes in the surrounding tissues. What is the minimum vessel size for wall nourishment via vasa vasora? Our guess is 200 μ, but the point ought to be documented.

There is an interesting difference in vasomotor responses between "red" and "white" muscles, the "red" being those rich in myoglobin and other iron containing respiratory pigments and which are primarily involved in long maintained contractions associated with the holding of a given posture against gravity. The "white" muscles, on the other hand develop tension for short periods of time in association with quick movements. Folkow and Halicka (65) studied blood flow, oxygen consumption, and average capillary filtration coefficients (an approximate indication of the number of open capillaries) in two kinds of muscle. Compared with white muscle, red muscle showed approximately two fold higher resting blood flows and double the maxima of blood flow and capillary filtration coefficient during muscle activity. Both kinds of muscle interfered with blood flow during tetanic stimulation, by mechanical compressive effects on the enclosed vasculature. Blood flow to red muscle supplied oxygen at rates in excess of the oxygen consumption rate; on the other hand, white muscle shifted to anaerobic metabolism at low stimulation rates and could not function other than anaerobically during tetanic stimulation. Red muscle was less susceptible to vasoconstrictor effects of sympathetic nerve stimulation; the maximal effect in red muscle was a 1.5 fold increase above resting resistance, whereas the maximal effect in white muscle was a 6 fold increase. Thus skeletal muscle cannot be regarded as uniform in response to vasoactive stimuli, for there are some muscles which are primarily

red, some which are primarily white, and many which are mixtures of the two kinds of fibres. These interesting differences in the microvasculature of the two kinds of muscle suggest a possibly useful experimental preparation for long term changes in muscle vascularity. Red muscles became white, and vice versa, when their respective motor nerves were reversed by sectioning and appropriate surgical anastomosis (66, 67). If there proved to be a large increase in vascularity in the white-becoming-red muscle, and a large decrease in the opposite case, it might turn out to be a convenient model system for study of long term vascularity changes.

Another important feature of skeletal muscle's vasomotion is the interaction between exercise hyperemia and sympathetic vasoconstriction. The studies of Remensnyder et al. (61) and of Kjellmer (62) show that when sympathetic vasoconstrictor fibres were stimulated during exercise, muscle blood flow was reduced only transiently. Thus, local 'follower' response of active hyperemia can over-ride an inappropriate centrally controlled command for constriction. A noteworthy aspect of Kjellmer's work (62), however, was that the reduction in vascular volume effected by sympathetic stimulation was not impaired during exercise. Similarly, an earlier study by Lewis and Mellander (68) had shown that, while arteriolar constrictive effects of sympathetic stimulation disappeared within 20-30 minutes during hypotensive perfusion, the sympathetically effected reduction in venous volume persisted. This reduction in vascular volume is almost altogether a venous phenomenon, and contributes, as shown by Guyton's studies on the pressure-flow relations of the whole vascular tree (31), significantly to the rise in cardiac output in exercise.

Thus there are three major regulatory characteristics of the vascular tree. They are:

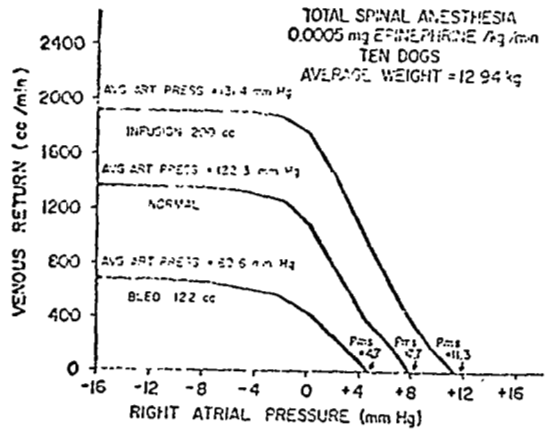
(1) Sympathetically mediated arteriolar constriction
(2) Metabolically dependent arteriolar and (?) venular dilatation
(3) Sympathetically mediated reduction in venous volume

These may be thought of as being superimposed upon the basic physiologically adaptive physical properties of the vascular tree. The basic physical properties, which have already been considered, can be summarized as a low capacitance arterial tree terminating in a high resistance that is followed by a high capacitance, low resistance venous tree that empties into the right atrium. Under a wide variety of conditions, right atrial pressure is within 1-2 cm H_2O of atmospheric, so the outlet of the venous tree may appropriately be said to be grounded.[1]

Guyton and his colleagues have devoted much experimental effort toward working out isovolumic outlet pressure-flow relations in the vascular tree. The empirical results are as shown in the next figure. It is essential to

[1] A major issue in CV system modelling may be contained in the question, "What determines the grounding of the outlet of the venous tree?"

have a clear view of how these data have been obtained.



Systemic vascular function curves in spinally anesthetized dogs infused, at a constant rate, with epinephrine. The three curves were determined after blood volume changes indicated. The rationale for total spinal anesthesia plus epinephrine infusion was to block autonomically mediated changes in vascular parameters in response to hemorrhage or transfusion and to variations in cardiac output (venous return). (31).

Guyton's technique (cf.(31))was a bit cumbersome, but it amounted to the interposition of a servo-controlled pump between the right atrium and pulmonary artery. We might note parenthetically that the essentials of the experiment are repeated daily in most major medical centers when cardiac patients undergo open-heart surgery using cardio-pulmonary bypass with a servo-controlled mechanical pump and oxygenator. In either case, there is independent control over the venous outlet pressure, while the servo-controlled pump delivers blood to the pulmonary artery (in Guyton's design) or to the systemic arterial tree (in total cardio-pulmonary bypass)in rates equal to the return of venous blood to the pump (venous return). It is now a part of everyday experience that venous return plateaus below a slightly subatmospheric outlet pressure (i.e. a suction), and falls quite precipitously with small increases in outlet pressure above atmospheric, as the figure shows. The plateau occurs as the thin-walled great veins alternately collapse and open, and do not transmit the suction back into the vascular tree. The pressure at zero flow has had a variety of names as it has cycled four times in the last century between discovery and oblivion in the hands, successively, of the Weber brothers, Ernest Starling, Isaac Starr and most recently, Arthur Guyton. Both the original and current term is "mean systemic pressure". This pressure may be viewed as arising from the interaction between the blood volume, the unstressed volume of the vascular tree, and the overall capacitance of the tree at volumes

in excess of the unstressed volume. Figures representative of existing conditions might be blood volume of 5000 cc, an unstressed volume of 4200 cc, an overall capacitance of 100 cc/mmHg, and a mean systemic pressure of 8 mmHg. It is important to note that the values of unstressed volume and overall vascular capacitance (which are extrapolations (31) of canine data to humans) come from measurements of blood volume and of mean systemic pressure. The physiological significance of mean systemic pressure lies only in its utility as an indirect measure of unstressed volume and vascular capacitance. In the view being advanced here, Guyton has obscured the volume and capacitance issues by regarding mean systemic pressure as a parameter of the operating cardiovascular system. There is, as Starling pointed out in his Arris and Gale lecture (69), some locus in the venous tree of the operating cardiovascular system upon which the mean systemic pressure is found. One can go further and compute the resistance to flow between that locus and the right atrium. But the locus upon which the mean systemic pressure exists under non-zero flow conditions, and the resistance to flow between that locus and the right atrium have no special physical or physiological significance, and to offer that resistance as a theory for the slope of the curve in the venous return figure, as Guyton does, only obscures the issue.

The slope of the curve in that figure arises from the following considerations. The blood volume, V_b is distributed as follows:

$$V_b = V_{pa} + V_{pv} + V_{sa} + V_{sv} + V_{ra} \qquad \text{Eq. 1}$$

where a signifies artery; c, capillary; v, vein; p, pulmonary; s, systemic. Thus, we are apportioning the system into the following lumps: pulmonary artery, pulmonary vein (here taken to include pulmonary capillaries and left atrium), systemic arteries, systemic capillaries, systemic veins, and right atrium (here taken to include the great veins within the thorax). Eq. 1 can be expanded to:

$$\begin{aligned} V_b = {} & V_{pao} + C_{pa}\,(P_{pa} - IPP) + V_{pvo} + C_{pv}\,(P_{pv} - IPP) + V_{sao} \\ & + C_{sa}P_{sa} + V_{sco} + C_{sc}P_{sc} + V_{svo} + C_{sv}P_{sv} + V_{rao} + C_{ra}\,(P_{ra} - IPP) \end{aligned} \qquad \text{Eq. 1'}$$

where C signifies capacitance; IPP, intrapleural pressure; P, pressure relative to atmospheric. For each blood pressure, other than right atrial, we can substitute the product of flow, Q, and the total downstream resistance, from the point in question to the outlet (which has a pressure P_{pv} for the pulmonary bed, and P_{ra} for the systemic) since the flows through each lump are all equal to one another and to cardiac output or venous return. Thus,

$$\begin{aligned} V_b = {} & V_{pao} + V_{pvo} + V_{sao} + V_{sco} + V_{svo} + V_{rao} + C_{pa}\,(P_{pv} - IPP) \\ & + QR_p + C_{pv}\,(P_{pv} - IPP) + C_{sa}\;P_{ra} + Q\,(R_{sa} + R_{ssv} + R_{slv}) \\ & + C_{sc}\;P_{ra} + Q\,(R_{ssv} + R_{slv}) + C_{sv}\;P_{ra} + QR_{slv} + C_{ra}\,(P_{ra} - IPP) \end{aligned} \qquad \text{Eq. 2}$$

where R_{ssv} is the resistance afforded by the small systemic veins or venules, and R_{slv} is the resistance afforded by the systemic large veins.

Solving for Q,

$$Q = \frac{\left[V_b - \text{unstressed volumes} - (C_{pa} + C_{pv} - C_{ra}) * (P_{pv} - IPP) - (C_{ra} + C_{sa} + C_{sc} + C_{sv}) P_{ra} + C_{ra} P_{pv} \right]}{R_{slv} (C_{sa} + C_{sc} + C_{sv}) + R_{ssv} (C_{sa} + C_{sc}) + R_{sa} C_{sa} + R_p Cpa} \qquad \text{Eq. 3}$$

Thus the slope of the relation depends upon the resistances around the circuit, each systemic resistance being weighted by the total upstream capacitance. Changes in unstressed volume, on the other hand have a negative influence on the zero flow pressure intercept of the curve, but not its slope; changes in blood volume have a positive effect on the zero flow pressure intercept, but not its slope. Of course, the assertions that neither blood volume nor unstressed volume changes influence the slope presumes that these volume changes can occur without some concomitant influence in any of the resistances. In the limit, of course, it is impossible to have the dimensional changes in the vascular tree that a change in blood volume would necessitate without some attendant resistance change. However, to a first approximation, it appears from Guyton's experimental work (cf. 31)) that blood volume or unstressed volume changes occur with minimal influence on the resistances around the circuit. Eq. 3, then stands as a more appropriate description of the outlet pressure-flow of the vascular tree than is the treatment offered by Guyton (cf. 31)), which carries mean systemic pressure as a system parameter.

One of the features of the vascular tree that Eq. 3 serves to emphasize, as did Guyton's analysis (31) and Grodin's earlier one (70) is the major role which the venous tree plays in forcing flow changes. (For this discussion, it suffices to note that the normal intact system will operate with right atrial pressure of approximately zero and a left atrial pressure (P_{py} in Eq. 3) of about 5 mmHg. If these values are assumed for the moment, we are spared the necessity of having to deal with the right and left heart equations to solve for cardiac output.) It is moderately well documented that the venous tree undergoes sympathetically mediated changes in the transition from rest to exercise. The question is whether these changes are primarily ones in unstressed volume or in capacitance. Capacitance has been the traditional term (cf.(71)); the term unstressed volume is here offered for the first time, insofar as we are aware, explicitly as a major vascular parameter. In evaluating these two possible parameters, we ought to consider the kind of evidence that is available for some kind of change in the veins during exercise. The recent review by Shepherd (71) surveys the literature and points out that two methods have been used to study vasomotor phenomena: measurement of pressure changes in a segment of vein in which both inflow and outflow have been arrested; measurement of volume changes in a limb in

response to varying venous pressures. Both methods show qualitatively that in exercise something happens to the veins that could be described either as a decreased capacitance or a decreased unstressed volume. We judge the parameter unstressed volume to be preferable for several reasons. First of all, in the areflex system, as Guyton has shown (cf.(31)), mean systemic pressure, and thus cardiac output, falls to zero with the loss of only about 15% of the blood volume. With reflexes intact, thus making possible venous changes of one kind or the other, a 25% blood loss (i.e., well in excess of the initial unstressed volume) diminishes cardiac output but does not reduce it to zero. How did compensation occur? It could not occur purely through a capacitance change, because once the system is bled down to below its initial unstressed volume, no amount of change in the extensibility of the wall will bring the numerator in Eq. 3 to positive values, or, to put it in other terms, no amount of change in the extensibility characteristics of venous walls will suffice even partly to restore intravascular pressures in underfilled vessels. Another difficulty with capacitance as a parameter, is that it is not at all clear that sympathetic stimulation reduces vascular capacitance, where capacitance is defined rigorously as the slope (dv/dp) of the pressure-volume curve. With sympathetic stimulation, the slopes of various vascular pressure-volume curves change erratically (50), whereas pressure-specific volumes seemingly invariably fall. For these reasons, we adopt the various unstressed volumes of the arbitrarily segmented vascular tree as a set of parameters which are part of the causality in the regulation of cardiac output.

The explicit modelling of cardiac output begins, then, with Eq. 3, or a set of underlying ohmic and pressure, volume, capacitance relations for each segment, together with two simplified one-parameter ventricular models, as we have advocated earlier. These may be regarded collectively as the unregulated system, awaiting causal links between system variables - pressures and flow (and their consequences) - and system parameters. The parameters that are subject to major physiological variation are, in one sense, relatively few: the conditional peak flows of each ventricle (which in the normal system would change nearly together and so might as a first approximation be treated as one), the unstressed volume of the systemic veins, the systemic arteriolar resistance, perhaps as well the resistance in the small systemic veins (which may show a metabolically linked vasodilatation), the pulmonary resistance, and blood volume. To date, one of us (JU), together with Dr. Donald A. Gall and Mr. Gary Barr have developed a steady-state, algebraic model for the unregulated system. With the extensive hybrid facilities soon to be available to two of us at the University of Southern California, we hope to incorporate the two major control loops of the metabolic vasodilatation and the baroreceptor reflexes, and also to begin exploration of plausible structures for our postulated exercise control algorithm.

At a more fundamental level, however, it is evident that we have no theory with which to account for the various dimensions of the vascular tree, nor the material properties of its walls. The best we have at present are some semiquantitative, moderately well documented ideas for the basis of physiological changes in these dimensions and material properties about a

central operating point. For this reason, we have referred to the aforegoing discussion as 'engineering' views of the cardiovascular system.

(c) A preliminary synthesis of the 'engineering' views of cardiac and vascular functions.- Consider the system's response to moderate levels of exercise, in which the cardiac output might rise about 2-1/2 fold, with a modest rise in arterial pressure, and little or no change in heart volume. The following parameter changes will acceptably well simulate these changes:

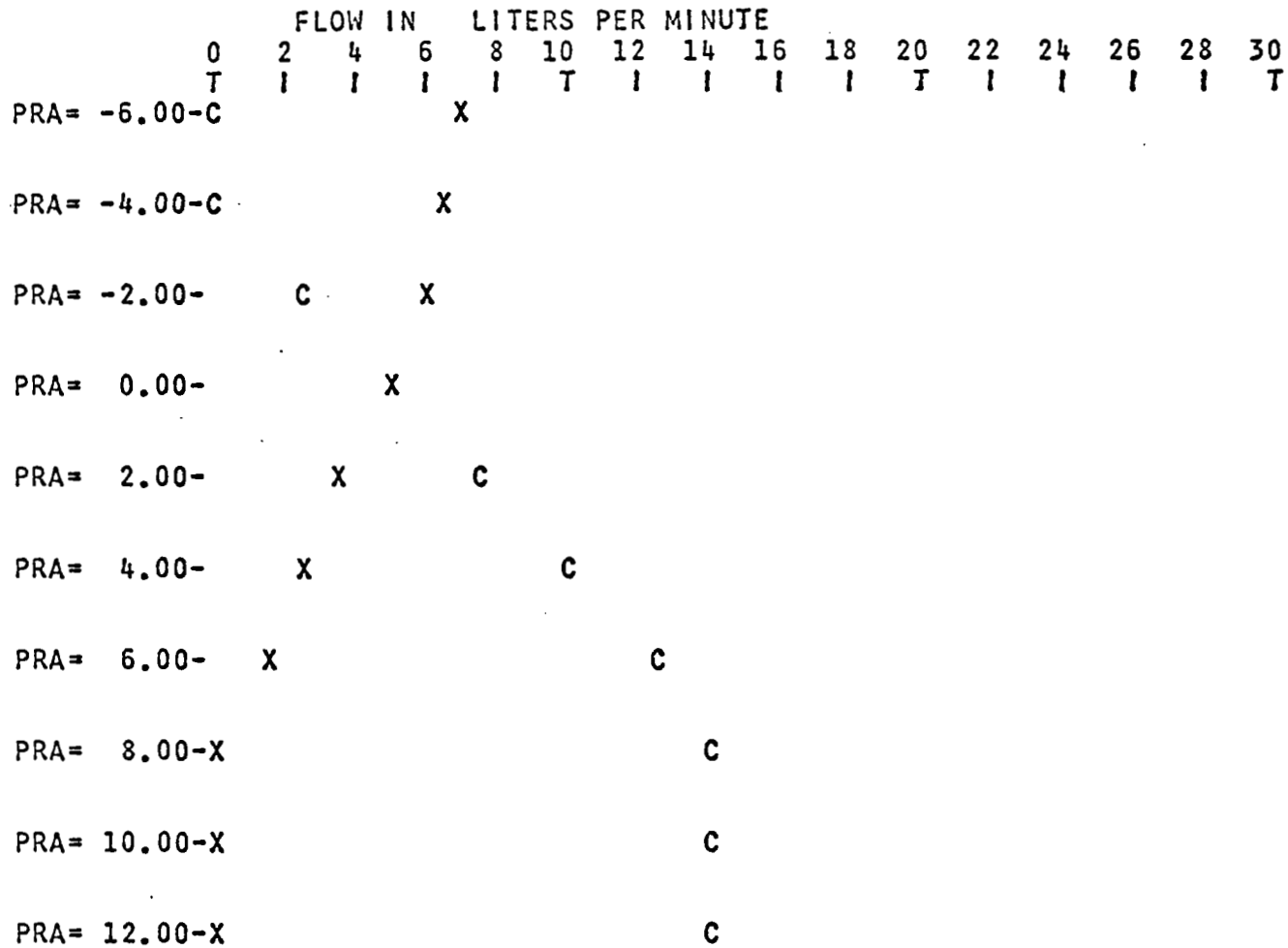
	rest	exercise
CPF(LH)	14.0	28
CPF(RH)	14.55	29
R_{sa}	14	6
R_{ssv}	4	1
R_p	2	0.5
V_{svo}	2	1.3

With these parameter alterations, the principal system variables undergo the following changes, as computed with our steady-state algebraic model.

	rest	exercise
Q_b	5	13.6
P_{sa}	95	110
P_{ra}	0	1.5
P_{sc}	25.1	28.6
P_{ssv}	5.0	15.0
P_{pa}	15.1	14.0
V_{ra}	0.56	0.59
V_{pv}	0.76	0.78
P_{pv}	5.0	7.2

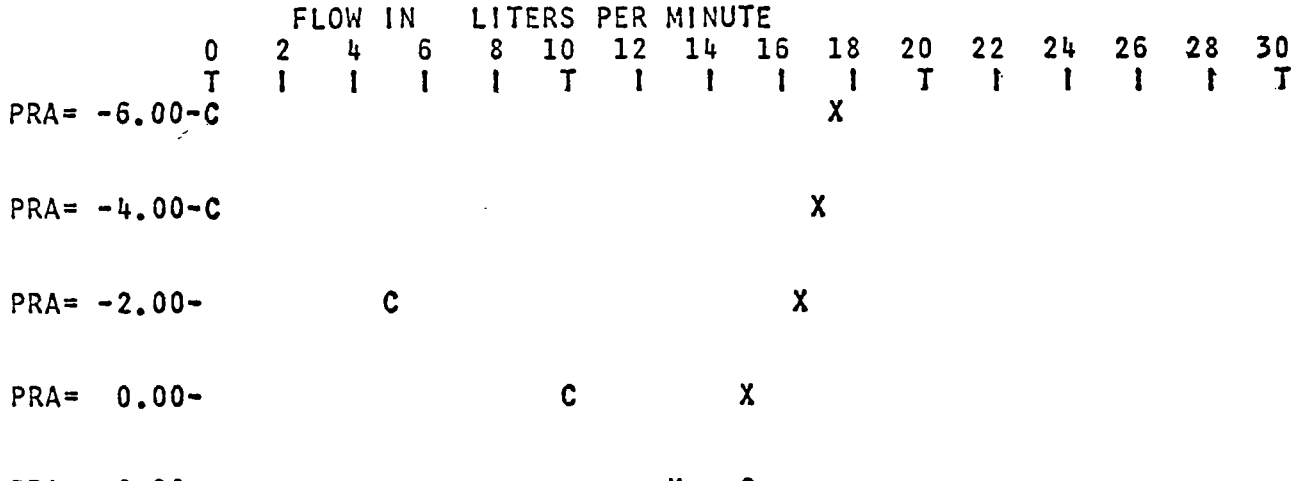
The following two computer 'figures' show the cardiac and vascular function curves computed, respectively, from the "rest" and "exercise" parameters. The computed curves conform to those presented by Guyton in his synthesis of the system (31), but here are explicable in terms of a physical model with the parameters we have discussed above.

It remains for the future to overlay baroreceptor reflex characteristics and metabolic hyperemic effects on this model, and then to infer a structure for the exercise algorithm that will force the regulated model into an operation that simulates the real system.



PLOT OF VENOUS RETURN (X) AND CARDIAC OUTPUT (C) IN LITERS PER MINUTE AS A FUNCTION OF RIGHT ATRIAL PRESSURE (PRA) IN mmHg

Systemic vascular function curve (X) and cardiac function curve (C) for the parameter values at rest, given in the text. The intersection of the two curves is the computed cardiac output and right atrial pressure. The values of other system variables of interest are given in the text.



PLOT OF VENOUS RETURN (X) AND CARDIAC OUTPUT (C) IN LITERS PER MINUTE AS A FUNCTION OF RIGHT ATRIAL PRESSURE (PRA) IN mmHg

Systemic vascular function curve (X) and cardiac function curve (C) for the parameter values in the text for exercise. The intersection of the two curves is the computed cardiac output and right atrial pressure. The values of other system variables of interest are given in the text.

(d) Some closing thoughts about describing the CV system.- For the development that we have outlined here, it has not been possible to arrive at a complete concurrence in opinion as to the best choice of cardiovascular systems parameters and systems relations - any more than it has among other students of the cardiovascular system. The difficulty with the parameters such as unstressed vascular volumes, conditional peak flows, or pressures that obtain under zero flow conditions is that they are not fully operational parameters, but are, in a sense, fictitious ones. They can only be defined under an extreme condition, and have to be linearly interpolated back to operating conditions compatible with the preservation of the system.

To accomplish one very minimal task and lay a foundation for future work, it is desirable to propose a new notation for cardiovascular, and probably other biological systems' parameters, in order to capture their obvious adaptive character. Three components of a system's parameter suggest themselves, denoted by the following subscripts: 00, to denote genetically determined, species specific values; 0, to denote an epigenetically arrived at status, as a result of long term adaptive changes; 1, to denote short term increments. Dynamic equations would usually refer to changes in the "1" components.

It has been the convention in descriptions of the cardiovascular system to begin, in the "1" time scale, with pressure-flow relations, the conservation of blood in the segments of the vascular tree, and so forth, just as we have done in the previous section. It is probably true that this approach will, one way or another, inevitably require extreme-state parameters. The alternative is to begin with the 00 time scale and to write the apparent algorithms of development. Examples of the form these algorithms might take stem from the following considerations. There is an apparent tissue demand for blood: approximately 500 to 1000 ml/kg/min. (setting aside second-round consideration of differences in various activity states, and specific tissue exceptions. One should note however, that as a result of the body's active control, it averages more nearly a demand of 100 ml/kg/min. at rest, or 300 ml/kg/min. at peak activity). There is an apparent blood volume law, approximately 70 ml/kg. Arteries and veins are made of materials with a given set of elastic-plastic properties, which set upper limits for pressures compatible with sustained vessel wall integrity. Out of these considerations, plus maybe others, we can dimly see emerging the constraints on vascular architecture. Similar considerations might lead to a developmental view of cardiac structure and range of function.

This view may be criticized as a romantic attempt to second-guess the logic of an unknown or non-existent designer. If one takes the view that the capacity for self-organization is a fundamental biological attribute, that it follows lawful relations, and that those lawful relations are not overly large in number, then a task is defined to model that process. The parameters of the models' equations will hopefully be more satisfactory than the extreme state ones we have discussed earlier. That is, however, a task for the future.

7. A Preliminary View of the Microcirculation in Skeletal Muscle.

While the survey of Iberall (72) provides an average view of the microcir-

culation, the studies of Smaje, Zweifach and Intaglietta on the microvasculature of the rat cremaster muscle (73) perhaps offers some more details of the vascular bed in skeletal muscles. Capillaries were noted to be oriented parallel to muscle fibres, and were 600 $\pm$ 200 μ long. About every 200 μ, a short cross connection was noted to an adjacent capillary. Intercapillary distance was 34 $\pm$ 2 μ, muscle cells being 20-25 μ in diameter. The capillary density was 1300/mm^2. The internal diameter was least at the arteriolar end of the capillary, and was about 2 μ smaller than the 7.5 μ erythrocyte diameter, thus strengthening Wiedeman's (74) observation in the bat wing that erythrocyte diameter exceeded capillary diameter. At the venular end, capillary diameters were about 1 μ larger than at the arteriolar end. Red cell velocity was about 700 μ/sec. Thus the average transit time for an erythrocyte was about 0.9 sec.

It is important to note that the cross branching from capillary to capillary occurred between capillaries supplied by the same arteriolar source and at essentially equal distances from that source. There is nothing to suggest that a given capillary might at its arteriolar end, receive blood from or contribute blood to, another capillary near its venular end. Nor was there evidence for other than cocurrent flow in adjacent capillaries. This work confirms the geometry that Krogh assumed many years ago upon which he and Erlang modelled oxygen transfer from blood to tissue cells (75).

Recent studies by Honig, et al. (76) on microvascular function in rat gracilis muscle is in essential agreement with the work of Smaje, et al., although they reported several quantitative differences. Most capillaries were 400-500 μ long with cross branches. Although this muscle has an appreciable third dimension, unlike the cremaster, the authors did not investigate the possibility of three dimensional symmetry in capillary distribution. The capillary density reported was considerably less than that reported in cremaster: 250/mm^2 vs 1300. Asphyxia increase capillary density 2.6 fold. They also showed cyclic fluctuations, with 2-4 minute period, in capillary density. The average capillary density was unaffected by denervation, adding support to the prevailing notion that capillary flow is under local, metabolic control.

Another noteworthy feature of the recent work by Zweifach and his collaborators is their attempt at refinement of the Starling hypothesis of transcapillary fluid exchange. Their work shows about a 10 cm H_2O pressure drop from the arteriolar to the venular end of the capillary. It also contributes some further information about the balance of forces between average capillary hydrostatic pressure, and colloid osmotic pressure of plasma. Finally, they showed, as Wiederhelm had earlier (77), that there was significant fluid exchange occurring in the post-capillary venule. Indeed, in their calculations, there was about 250 cm^2 of capillary wall per cc of muscle, plus an additional 270 cm^2 of venular wall per cc of muscle, which also permits fluid exchange, although they have not quantitated it with an explicit filtration coefficient as they have the capillaries themselves.

Thus the study by Smaje, et al., provides one additional attempt at a quantitative view of the architecture, dimensions, and fluid permeability

characteristics of a particular muscle bed. It happens, of course, to be a very thin sheet of muscle - essentially a two-dimensional system. It leaves open the nature of the cross connections between adjacent capillaries in the much more common case of three-dimensional muscle capillary beds. The work by Smaje, et al., nicely complements the 1967 review by Forster (78) on oxygen exchange between capillaries and muscle cells, at which time the critical dimensions and red cell transit times in functioning muscle were still uncertain because of the possibility of ante-mortem or fixation artefacts. At this time, one can only hope that the improvements in technology in microvascular research, plus attention to important quantitative details, will provide a better documented model of respiratory gas, water, and solute exchange at the microvascular level, which will supplement the views that we have today of these characteristics in overall vascular beds (79).

Two specific examples of very recent work will serve to illustrate future problems. Berne and his colleagues reported at the 1970 FASEB meeting that they had found oxygen tensions in arterioles far below simultaneous arterial levels (80). These measurements were made with the Whelan oxygen microelectrode. These observations were completely unexpected, since it has always been regarded as axiomatic that transfers between blood and tissues occur in any quantity only when there is but a single endothelial layer separating blood and tissue. This axiom is doubtless true for water and water soluble solutes. In retrospect, perhaps it is not so surprising that a lipid soluble substance like oxygen might diffuse readily across a multicellular arteriolar wall. In any case, it becomes a more difficult problem to assess the relative amounts of oxygen exchange between the red cell and the skeletal muscle, as the red cell makes its passage through the microvasculature. Whether the major exchange takes place in the capillary, or is facilitated by the mechanical deformations which result from single file flow of erythrocytes down a capillary of less-than erythrocyte diameter, or whether the arterioles contribute significantly to tissue oxygenation now requires further study.

Finally, Cardon et al., (14) have measured "spontaneous" fluctuations in erythrocyte flow in the thin layer of skeletal muscle that lies beneath the skin in the dorsum of the mouse. Marked variations in erythrocyte flow, with three discernible, superimposed periods of 30, 100 and 320 seconds were observed; these temporal patterns were altered by hypoxia. It is puzzling that these investigations observed no corresponding dimensional changes in capillary diameter, in precapillary sphincteric activity or in arteriolar diameter. These observations leave entirely open the question of the gating mechanism. It is tempting to speculate that these "spontaneous" fluctuations in erythrocyte flow in single capillaries are the substratum of the "spontaneous" fluctuations in whole-body oxygen consumption (18).

Bibliography

1. Altman, P., Dittmar, D., (Ed.), "Biology Data Book", FASEB, Washington, D. C., 1964.
2. Patel, D., DeFreitas, F., Greenfield, J., Fry, D., J. App. Physiol. 18, 1111, 1963.
3. White, L., Haines, H., Adams, T., Comp. Biochem. Physiol. 27, 559, 1968.
4. Patterson, J., Goetz, R., Doyle, J., Ann. N. Y. Acad. Sci. 127, 393, 1965.
5. Bartels, H., Amer. Heart J., 72, 1, 1968.
6. Costill, D., J. Appl. Physiol. 28, (3), 251, 1970.
7. Mayer, J., Thomas, D., Science 156, 328, 1967.
8. Parks, J., Am. J. Physiol. 216, 666, 1969.
9. Terman, M., Terman J., Science 168, 1242, 1970.
10. Wolf, S., Internal. J. Psychobio. 1, 27, 1970.
11. Sargent, R. M., Proc. Roy. Soc. B., 100:10, 1926.
12. DeLanne, R., Barnes, J. R., Brouha, L., and Massart, F., J. Appl. Physiol. 14:328, 1959.
13. Iberall, A. S., Cardon, S. Z., NASA CR-219, May, 1965.
    Iberall, A. S., J. Basic Eng'g 81:96, 1960; 81:103, 1960; 82:513, 1960.
14. Cardon, S. Z., Ostermeyer, C. F., and Bloch, E. H., Microvasc. Res. 2:67, 1970.
15. Iberall, A. S., Ehrenberg, M., Cardon, S. Z., Simenhoff, M., Metabolism 17:1119, 1968.
16. Benzinger, T. H. and Kitzinger, C., in "Temperature, Its Measurement and Control in Science and Industry", Vol. 3, Part 3, New York: Reinhold, 1963.
17. Iberall, A. S., Cardon, S. Z. and Jayne, T., NASA CR-141, Jan. 1965.
18. Goodman, L., IEEE Trans. Biomed. Eng'g BME-11:82, 1964.
19. Cowan, C. R. and Solandt, B. M., J. Physiol. 89:462, 1937.
20. Oswald, I., Merrington J., Lewis, H., Nature, 225:959, 1970.
21. Hellman, L., Nakada F., Curti J., Weitzman, E., Kream, J., Roffwarg, H., Ellman, S., Fukushima , D., Gallagher, T., J. Clin. Endocr. 30:411, 1970.
22. Hellman, L., et al. J. Clin. Endocr. 30:686, 1970.
23. Heymans, C., and Neil, E., "Reflexogenic Areas of the Cardiovascular System", Boston: Little Brown, 1958.
24. Levison, W. H., Barnett, G. O., and Jackson, W. D., Circ. Res. 18:673, 1966.
25. Astrand, P. O., Caddy, T. C., Saltin, B., and Stenberg, J., J. Appl. Physiol. 19:268, 1964.
26. Bevegard, S., Holmgren and Jonsson, B., Acta Physiol. Scand. 49:279, 1960; Acta Physiol. Scand. 57:26, 1963.
27. Ekblom, B., Astrand, P. O., Saltin, B., Stenberg, J., and Wallstrom, B., J. Appl. Physiol. 24:518, 1968.
28. Douglas, F. G. V., and Becklake, M. R., J. Appl. Physiol. 25:600, 1968.
29. Saltin, B., Blomqvist, G. Mitchell, J. H., Johnson, R. L., Jr., Wildenthal, K., Chapman, C. B., Circulation, Suppl. VII to vols XXXVII and XXXVIII: VII-I, 1968.

30. Forsyth, R., Nies, A., Wyler, F., Neutze, J., Melman, K., J. Appl. Physiol. 25:736, 1968.
31. Guyton, A. C.,"Cardiac Output and its Regulation" Philadelphia. Saunders, 1963.
32. Cunningham, D. J. C., Circ. Res. Suppl I to vols. XX and XXI: I - 22, 1967.
33. Dejours, P. in D.Cunningham, B. Lloyd "The Regulation of Human Respiration", Blackwell, 1963.
34. Clarke, N. P., Smith, O. A., and Shearn, D. W., Am. J. Physiol. 214:122, 1968.
35. DiCara, L., and Miller, N., Science 159:1485, 1968.
36. Asmussen, E., Circ. Res. Suppl. I to vols XX and XXI: I-2, 1967.
37. Spotnitz, H. M., Sonnenblick, E. H., and Spiro, D., Circ. Res. 18:49, 1966.
38. Starling, E. H.,"Linacre Lecture on the Law of the Heart". London: Longmans, Green, 1918.
39. Sarnoff, S. J., and Mitchell, J. H., in "Handbook of Physiology",Vol. 1, Section 2, Hamilton, W. F., and Dow, P., (Eds.). Washington: American Physiol. Society, 1962, p. 489.
40. Braunwald, E., Ross, J. Jr., and Sonnenblick, E. H.,"Mechanisms of Contraction of the Normal and Failing Heart". Boston: Little Brown, 1967.
41. Sagawa, K., in "Physical Bases of Circulatory Transport" ed. by E. B. Reeve and A. C. Guyton. Philadelphia: Saunders, 1967, p. 141.
42. Herndon, C. W., and Sagawa, K., Am. J. Physiol. 217:65, 1969.
43. Stone, H. L., Bishop, V. S., and Guyton, A. C., Am. J. Physiol. 204:16, 1963.
44. Fenn, W. O., J. Physiol. 58:175, 1923.
45. Hill, A. V., Proc. Roy. Soc. B. 126:136, 1938.
46. Hill, A. V., Proc, Roy. Soc. B. 159:297, 1964.
47. Peachey, L. P. Ann. Rev. Physiol. 30:401, 1968.
48. Mellander, S., Acta Physiol. Scand. Suppl. 176:1, 1960.
49. Sarnoff, S. J., Gilmore, J. P., Brockman, S. K., Mitchell, J. H., and Linden, R. J., Circ. Res. 8:1123, 1960.
50. Alexander, R. S., Circ. Res. 2:405, 1954.
51. Salzman, E., Circ. Res. 5:149, 1957.
52. Scher, A. M. and Young, A. C., Circ. Res. 12:152, 1963.
53. Stegemann, J., and Muller-Butow, H., Arch. Ges. Physiol. 287:247, 1966.
54. Lamberti, J. J., Jr., Urquhart, J., and Siewers, R. D., Circ. Res. 23:415, 1968.
55. Koch, E.,"Die Reflektorische Selbststeuerung des Kreislaufes", Dresden: T. Steinkopff, 1931.
56. Sagawa, K., and Watanabe , K., Am. J. Physiol. 290:1278, 1965.
57. Allison, J. L., Sagawa, K., and Kumada, M., Am. J. Physiol. 217:1576, 1969.
58. Ead, H. W., Green, J. H., and Neil,E.,J. Physiol.118:509, 1952.
59. McCubbin, J. W., Green, J. H., and Page, I. H., Circ. Res. 4:205, 1956.
60. Kety, S. S., in"Handbook of Physiology", vol. III, section I, Washington: American Physiological Society, 1963, 1751.
61. Remensnyder, J. P., Mitchell, J. H., and Sarnoff, S. J., Circ. Res. 11:370, 1962.
62. Kjellmer, I., Acta Physiol. Scand. 63:450, 1965.

63. Skinner, N. S., Jr., and Powell, W. J., Circ. Res. Suppl. I to vols XX and XXI: I-59, 1967.
64. Skinner, N. S., Jr., and Costin, J. C., Fed. Proc. 29:116, 1970. (abstract).
65. Folkow, B., and Halicka, H. D., Microvasc. Res. 1:1, 1968.
66. Buller, A. J., Eccles, J. C., and Eccles, R. M., J. Physiol. 150:399 and 417, 1960.
67. Romanul, F. C. A., and Van Der Meulen, J. P., Arch. Neurol. 17:387,1967.
68. Lewis, D. H., and Mellander, S., Acta Physiol. Scand. 56:162, 1962.
69. Starling, E. H., Lancet 1:569 and 652, 1897.
70. Grodins, F. S., Quart Rev. Biol. 34:93, 1959.
71. Shepherd, J. T., Circ. Res. Suppl. I to vols XX, XXI; I-70, 1967.
72. Iberall, A., Math. Biosc., 1:375, 1967.
73. Smaje, L., Zweifach, B. W., and Intaglietta, M., Microvasc. Res. 2:96, 1970.
74. Weideman , M. P., Circ. Res. 12:375, 1963.
75. Krogh, A., J. Physiol. 52:409, 1919.
76. Honig, C. R., Frierson, J. L., and Patterson, J. L., Am. J. Physiol. 218:937, 1970.
77. Weiderhelm, C. A., in "Physical Bases of Circulatory Transport", ed., by B. A. Reeve and A. C. Guyton. Philadelphia, Saunders, 1967, 313.
78. Forster, R. E., Circ. Res., Suppl. I to vols XX and XXI:I-115, 1967.
79. Landis, E. M. and Pappenheimer, J. R.,"Handbook of Physiology", Sec. II, Vol. II. Washington:American Physiological Society, 1963, 961.
80. Duling, B. R., and Berne, R. M., Fed. Proc. 29:427, 1970 (abstract)

IX. A PHYSICAL VIEW OF METABOLIC SYSTEMS - THE BIOCHEMICAL TRANSITION FROM ORGAN SYSTEMS TO THE HIGHER CENTRAL NERVOUS SYSTEM

1. General Principles of Metabolic Systems

(a) Introduction. - The characterization of living systems has become the outstanding scientific challenge of the second half of the twentieth century. Organization of matter into hierarchical, growing, self-replicating systems intrigues philosophers, chemists, physicists, engineers, mathematicians and biologists, all of whom have brought to bear their methods, concepts and viewpoints. But in spite of the multiple approaches to the problem of the characterization of living systems, the suspicion remains that the essence has been lost, or missed. The great variety and rich detail of specialization among living forms present a confused landscape, and only a few, general principles concerning its origins and nature have so far emerged. These few principles include the following:

(1) Some neo-Darwinian view of the origin of the species - including random variation, genetic drift, competition and selection;

(2) A neo-Mendelian view of unit, discrete, genetic transmission of inherited characteristics;

(3) Recognition of the nucleic acid, polymeric basis of coding of the genotype;

(4) Recognition of the cell as the minimal unit with all the attributes that define a system as 'living' (although here some doubt exists as to whether the organelle, such as the mitochondrion, was a primordial, independent cell);

(5) Recognition of catalysis by functional proteins as the basis of extremely versatile chemical processing under conditions of limited terrestrial pressures and temperatures;

(6) Appreciation that the biosphere is a delicate system spread as a thin veneer on earth, sustained by mass and energy transfers involving the sun, microorganisms, plants and animals;

(7) Knowledge that the storage forms of the Gibbs free energy used in the support of synthesis, secretion, repair, maintenance and motion in biological systems are few in number, and adenosine triphosphate (ATP) is the archetype of these compounds;

(8) Conviction that biological systems do not disobey the known laws of thermodynamics, or physics and chemistry generally.

Although each item on the above list could be made the subject of an animated debate among biologists, probably all would agree that an awkward incompleteness haunts this list. The rudiments of reliable replication within a wide variety of environmental changes still elude us. The physical principles behind stability, information flow, low-error coding and the extraction of order (consisting of geometrical forms in space and function in time) from chaos are not specified. The issue is not so much whether a vitalistic principle exists, as whether modern physics is complete to the point of providing the basis for the description of life. In what follows it is assumed that to explain life something new may have to be added to the application of physics, as was the case when statistical mechanics followed classical thermodynamics, but that a new physics (in the sense that quantum mechanics and relativity were new compared to Newtonian physics) is not going to be required.

To set the scene for the preliminary description that follows, several of the guidelines employed should be made explicit. To begin, it is recognized that the design of biological systems is given us by illustration after illustration, and our task is to guess the blueprint. We are not free to invent life as it might have been designed. Therefore, we must start by considering the details of living systems as revealed by biological study. We are aware, however, that certain simplifying generalizations may have been missed by biologists, and that it is both desirable and necessary in the quest to see the landscape whole, to stand back from it and observe the gross features - the mountain ranges and valleys - not just the pebbles and grains of sand mined from experimental laboratories.

The thermodynamic questions are major ones to be answered. For example, over what time period, and in what manner is closure achieved with respect to the first law of thermodynamics applied to open systems exchanging both matter and energy with surrounding, environmental systems? In the answers we hope to find the physical constraints on the behavior of man - neglecting such poetic details as that some men may fall in love only with women with red hair. (Important as this local behavior is in the life of a particular individual, the machinery of life and the advance in numbers of the species does not involve redness of hair as a general principle - though assortive mating - of which the preference for red hair is merely an example - may itself have powerful global effects for a species.)

Living systems are systems, and they are hierarchical. The level in the hierarchy at which the main, thermodynamic features clearly emerge is higher than the molecular and we shall attempt to identify that level in this analysis. However, since few believe that complex form and function originated de novo, we start, as is customary, with a consideration of the lower levels of living systems. The main hierarchies of life are given below in Table 1.

Table 1

Hierarchies of Living Forms in Ascending Order of Complexity of Structure and Function

1. Inorganic materials
2. Organic monomers (purines, pyrimidines, amino acids, fatty acids, sugars)
3. Nucleic acid polymers as carriers of information
4. Protein polymers as carriers of control and function, and to some extent information (e.g., immunological reactions)
5. Viruses (nucleic acid-protein systems)
6. Biochemical chains - multienzyme sequences
7. Organelles as functional units for energy trapping and conversion, and unit processing

    a. Membranes
    b. Mitochondria and chloroplasts
    c. Ribosomes
    d. Lysosomes
    e. Vacuoles
    f. Golgi apparati
    g. Nuclei
    h. Nucleoli
    i. Centrosomes
    j. Kinetosomes (basal bodies)
    k. Pinocytotic vesicles

8. Cells-microorganisms
9. Slime molds - cellular aggregates
10. Simple multi-cellular animals and plants
11. Tissues and organs of complex animals
12. Organ systems in plants and animals
13. Complex animals and plants
14. Species of higher animals and plants
15. Ecosystems
16. Social systems
17. Terrestrial biosphere

In the above table, items 1 and 2 are studied by those interested in chemical evolution. Items 2-5 are studied mainly by molecular biologists. Items 3-7 are of interest to biophysicists; items 8 and 9 are the domain of the cellular biologist; items 10-14 interest the animal or plant physiologist. Items 15-17 constitute the now popular field of ecology.

In the present analysis the levels of greatest importance are 6-12. However, some consideration must also be given to macromolecular levels (3-5).

(b) Macromolecular biology. - The theoretical aspects of the evolution of biological macromolecules have been developed by numerous men, most recently Manfred Eigen, whose ideas are paraphrased here.

Subcellular biology is characterized by coupled catalytic chains and feedback at the lowest level of stable function. Because of feedback, effects are not sums of independent causes and both proteins and nucleic acids are necessary simultaneously for matter to replicate. (The 'chicken or egg' question is not meaningful when applied to the origin of biological macromolecules, because they are part of a feedback system and have not evolved separately and independently.) Replication is a system attribute and not a molecular attribute. The processes involved at this low level in biology are not random, and computations on the probability of origin of life based on an assumption of randomness are both naive and fallacious. Self-replication is a non-random, cybernetic process requiring both proteins and nucleic acids for it to exist at all. The evolutionary question of importance, then, is how did the cybernetic system of coupled nucleic acid-protein polymers arise, and what principles guided its evolution?

The complementarity of nucleic acids, which is inherent in their structure will lead to replication in the absence of proteins, but the error rate is very high, and the replication rate is very low. In the presence of protein catalysts, however, the error rate is low and the replication rate can be very high. The functional attributes of proteins that cause these effects reside at the tertiary level of structure, that is, in the folding of partially helical, covalent sequences of amino acids.

The smallest proteins have about 100 amino acids in the primary covalent sequence, and there are twenty different common amino acids available in nature. The number of possible primary chains in proteins is therefore 20^{100} (approximately 10^{130}). Of this overwhelmingly large number of possible proteins, only a small number actually exist, and of these so far only three have been analyzed completely for their tertiary structure by X-ray crystallography (myoglobin, hemoglobin and chymotrypsin). Ribonuclease is a fourth protein whose structure-function relationship is well understood.

(c) Quaternary protein structure and reaction control. - At the quaternary level of protein structure, the folded chains spontaneously aggregate, and several energy minima appear. Conformational changes from one to the other then become possible, with functional conversion as a result. Theories of functional conversions of proteins, that is 'conformational activation', have been advanced by Koshland and Monod, among others. Three mechanisms have been proposed to explain the functional changes in state of single proteins:

(1) Induced fit (Koshland). - In this view small molecules or large in the proximity of a protein induce by chemical forces a change in the conformation of the protein that then reacts with them.

(2) Allosterism (Monod). - In this view the binding of a small molecule at a site on a large one causes a conformational change that affects an active site elsewhere on the large molecule, changing its catalytic properties at that remote site.

(3) All or none. - In this view proteins act as logical switches, and are either in an active state, or in an inactive state, and are interconverted perhaps by the action of another catalyst (e.g., the activation of phosphorylase through its conversion from phosphorylase b to phosphorylase a by an enzyme).

In all the above cases the performance of the functional protein is amplifier-like, with small molecules providing a grid-control over the turnover number of the catalytic proteins. Marked amplification of a chemical signal can thus occur. Since the above mechanisms are not in any sense mutually exclusive, all may exist. The important point is, for present purposes, that the degree of function associated with at least some proteins can be varied in either a continuous or an all-or-none fashion. All three mechanisms provide a basis for positive or negative parametric feedback in biochemical chains, or coupling between them. The effects are parametric because they affect chemical affinities in catalyzed reactions. From such feedback chains, and the amplification characteristic provided by conformational activation of catalysts, biochemical control circuits can arise, and these may be characterized by oscillations.

From these considerations it appears that proteins at the level of tertiary and quaternary structure have two important properties: a) function (e.g., catalysis) and b) control. They do not have complementarity or replication. Such features are found only in nucleic acids. Once there is one protein, the chance for the appearance of another of the same amino acid sequence is zero. To create another similar protein it is therefore necessary to have an instruction.

The elemental, minimal aggregate of structure and function underlying life thus can be summarized as follows:

(1) Instructions (codes) lie in the relatively simple, complementary nucleic acid polymers.

(2) The instructions are translated into primary amino acid sequences of proteins.

(3) Primary amino acid sequences lead spontaneously to secondary, tertiary and quaternary structure of proteins.

(4) From tertiary and quaternary protein structure arise larger scale structure, function and control in chains, feedback loops, and organelles.

(5) From chains and loops arise higher level processes such as respiration, replication, and other system behavior characteristic of life, including adaptation and hereditary evolution.

(d) Evolution. - To describe the evolution or adaptation of the nucleic acid-protein system, some function, possibly new to thermodynamics, may be involved. In Eigen's view, it is a 'value' function. What seems to be needed

is a variant of cybernetics or game theory based on kinetics, irreversible thermodynamics and fluctuations. A binary information theory (communication theory, coding theory) is not appropriate to the problem of the adaptation or evolution of nucleic acid-protein systems. The issues are not binary, not 'all or none'. Instead they have to do with a 'value' based on the appearance or stabilization of a function in the face of fluctuations and selection forces. The selection principle, at the primordial level, is amplification based on speed of reaction; that is, the biochemical chain that produces the most functional proteins, with the largest turnover number, ultimately dominates the population of chains in a reaction mixture. The value function may have the nature of one or more weighted biochemical properties.[1] The 'negentropy' of Brillouin perhaps overemphasizes order in the physical dimensions of phase space. The value function in biology may have to emphasize order in some 'information space'. Stabilization at this lowest level of processing is achieved by feedback loops, and once a loop is formed, all its members are preserved in the face of fluctuations of the environment of the system.

The above principles of macromolecular action and evolution were recently validated in a brilliant experiment by Spiegelman. He demonstrated that a virus placed in a synthetic medium replicated in a very short time forms that were optimal for fast processing of the nutrients in the artificial medium. The replicating virus was then discovered to have lost all infectivity for the bacterial cells which originally provided the necessary environment for replication. In brief, the virus had 'adapted' or evolved to a new optimal performance characterized by the value evidently placed on fast processing.

In summary, many of the attributes so striking about living systems, including their replicability and adaptability are found at subcellular levels of organization. At these levels already the properties of control, communication, stability, and adaptive behavior are to be found. An essential feature of the organization of matter bestowing these properties on biochemical chains is amplification, positive and negative feedback, and coupling among reaction chains. This latter feature will be described in more detail below.

(e) Metabolic systems. - Metabolic systems are composed both of biochemical chains of catalysts with their associated nucleic acids, and of membrane transport systems.

Metabolic chains involve ten to 100 steps, each step catalyzed. The chain may or may not have feedback or feedforward, but it always has coupling nodes to other biochemical processes at one or more points. Feedback, feedforward, or coupling effects are exerted through several mechanisms: 1) through the amount of active catalyst at a rate-limiting step; 2) through the

[1] A potential introductory theory to a 'value' function was cast in the introductory Section I, in the form of summational invariants of the statistical mechanical ensemble that a set of organized cells in an organism are members of.

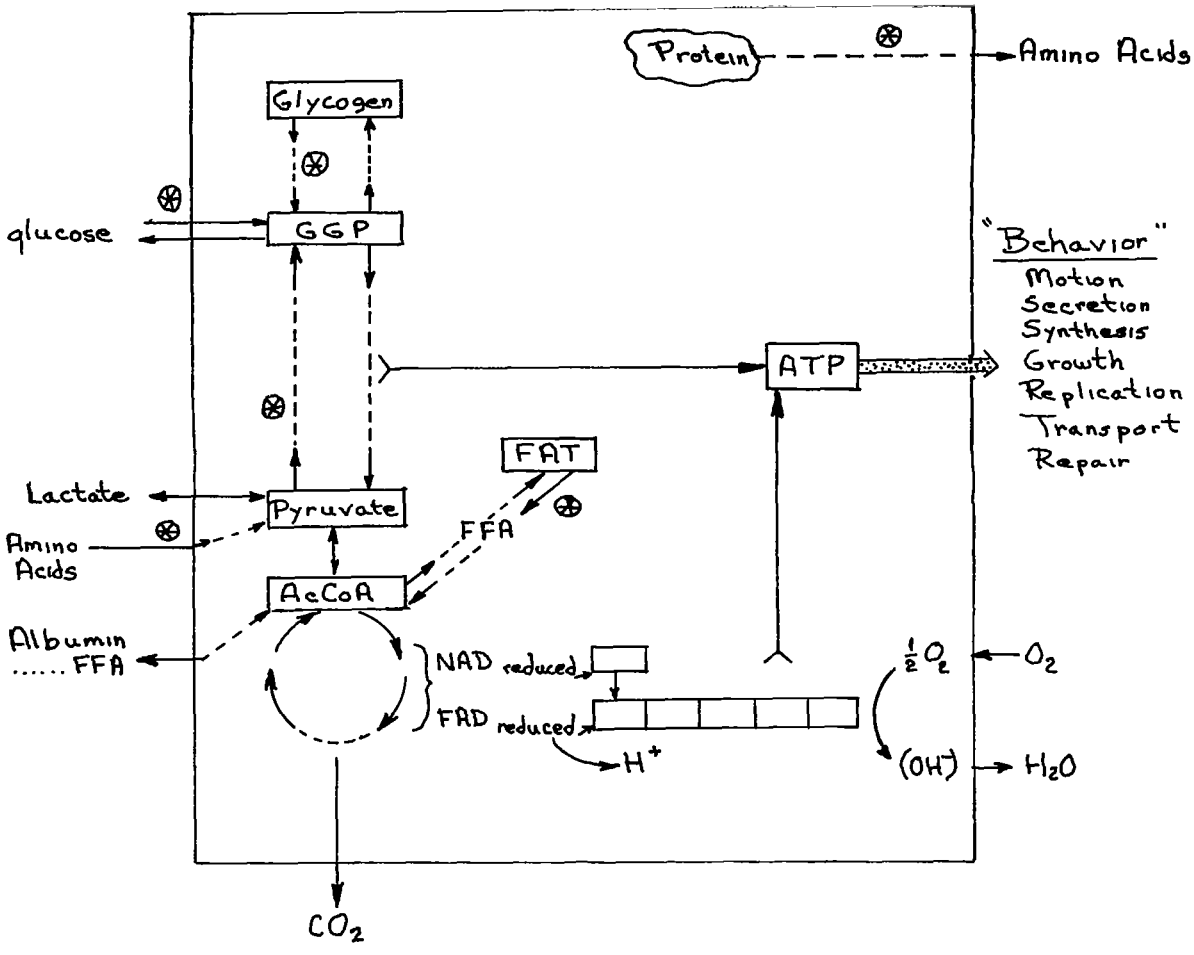
conformation of a given catalyst at a rate-limiting step or; 3) through the supply of substrate to a step in the chain.

(f) Control. - At least one kinetically irreversible step is usually present in a chain, with a different pathway involving different catalysts for the forward and backward pathways. Other steps in the metabolic chains are kinetically reversible and the forward and backward pathways are the same. Control is therefore exercised either at rate-limiting steps, or at kinetically irreversible steps. An important feature of metabolic chains is that a given chain may be subjected to very different parametric influences, depending upon which enzyme in the chain is rate-limiting. In different cells, the rate-limiting step of a chain may occur at different points in what otherwise would be the same chain. Operationally, the effect is profound and the metabolic characteristics of the two cell types may be rather different.

The control mechanisms that vary the amount of a given catalyst at a rate-limiting step must affect nucleic acid-dependent protein synthesis. To do so, chemical signals act either as inducers (derepressors) or repressors, and may be metabolites (i.e., substrates or products of reaction chains) or hormones. In any case, the activity of the genome is altered, and the rate of synthesis of the catalyst for the rate-limiting step in a metabolic chain is changed. Another control mechanism involves conformational activation of a catalyst at a rate-limiting step, and the chemical agents responsible for such activation are of three types: 1) metabolites (substrates, products, cofactors including such ions as potassium, magnesium, calcium, and such energy sources as ATP); 2) hormones; and 3) special signal molecules such as cyclic adenylic acid (cyclic AMP) that act as intracellular messengers. Hormones are chemical agents that affect process rates by acting as signal molecules transferred between cells, without themselves being a substrate, product or cofactor of the process within cells.

Control of the action of a metabolic chain achieved through variations in supply of substrate to a step in the chain involves orthodox mass action effects. Sharing of a common reactant by two or more chains is the most common basis of coupling in the metabolic system, and the effects are again those of mass action. If for example a chain produces ATP, but another reaction (ATP-ase) is present, the action of this second enzyme, outside the chain, can affect the rate of the processing in the whole chain because in the main path in the chain ATP levels are governing at one step. This example merely illustrates a common type of coupling among metabolic processes. The coupling effects can shift the rate-limiting point in a chain, and thus create mass action effects at branch points in chains that reroute molecular traffic through the freeway pattern of the metabolic chains. Extremely well known examples of chain interaction include the Pasteur and Crabtree effects. In the former, the consumption of oxygen by the electron transport chain produces effects that inhibit the production of energy through glycolysis; in the latter the reverse is true, and activation of glycolysis inhibits respiration.

(g) Occurrence of chains. - Certain metabolic chains are ubiquitous (glycolysis, photosynthesis, fatty acid oxidation, tricarboxylic acid cycle-electron transport) whereas others are specialized and occur only in certain cells (e.g., corticosteroido genesis in the adrenal gland). Nucleic acid-protein interactions, as previously stated, underlie all living systems as far as we know. The main emphasis in the present analysis shall concern those chains that are ubiquitous, and all of these appear to be involved with bioenergetics. They provide for the production and storage of Gibbs free energy in living systems. Although these chains are regarded as ubiquitous, it must be pointed out that in fact several different solutions have evolved for the energy trapping, production and storage processes required for life. For example, ATP production is carried out by microbial primitives through thioester production (lactyl glutathione); it is carried out by microbial specialists using amino acids, purines, choline, creatinine, and unusual sugars as carbon and energy sources. It is carried out by metabolic generalists, which are the organisms of interest in this analysis, through the processes of glycolysis, the Krebs cycle with oxidative phosphorylation (mitochondrial electron transport) or by photophosphorylation (photosynthesis). The following figure relates these metabolic chain processes in a composite animal cell. These are the chains responsible for the major thermodynamic engine characteristics of animal systems generally. Plant systems differ only in having an accessory means of trapping energy through photosynthesis. Otherwise the metabolic machinery within the cells is similar to that found in animals, and is shown in the figure.



Bioenergetics of a Composite Animal Cell, Showing Features Present in Liver, Muscle and Fat

Only major metabolic chains are shown. Control points are indicated by asterisks. These are points of action of insulin, growth hormone, epinephrine, glucagon or cortisol. Abbreviations used are:

ATP = adenosine tryphosphate;
NAD = nicotinamide adenine dinucleotide;
FAD = flavin adenine dinucleotide;
G6P = glucose-6-phosphate;
AcCoA = acetyl-coenzyme A;
FFA = free fatty acids.

(h) Membrane transports and gating. - A second component of metabolic systems is membrane transport. It can be seen from the figure that substrates must cross the phase boundary between cell environment and cell interior. Control by gating is possible at the boundary, and does indeed occur. The operation of the transport system at the boundary has profound effects through mass action on the substrate loading of the metabolic chains within the cell. Thus transport processes not only provide selectivity, but they can shift the operation from one chain to another, depending on which substrate is admitted. An interesting feature of transport phenomena is that they involve the coupling between a scalar exergonic process to a vectorial endogonic transport which is enzyme or carried-mediated.

An appreciable amount of control in biological systems is exerted at the level of transport through the limiting cellular membrane.

(i) Overall metabolic performance-metabolic states. - The overall performance of metabolic systems defines the physical chemical state of the living system. It determines the particular, gross mode of action of the biological system, including thermogenesis, motion, replication, secretion, repair, growth, differentiation, metamorphosis, flowering, etc. The extent to which it might determine the behavioral state of individual higher animals with elaborate nervous systems is given consideration at another point in this report. The number of distinct metabolic states that are recurrent is apparently not high. Below, an attempt is made to relate metabolic states to behavioral modes, but it must be admitted that behavioral modes are somewhat anthropomorphically defined.

Table 2

Recurrent Metabolic States

1. Metabolic States in Deciduous Plants

   a. Seasonal growth
   b. Seasonal flowering
   c. Seasonal abscission and dormancy
   d. Circadian photosynthesis and other rhythms

In the higher plants a behavioral mode is apparently synonomous with a metabolic state, and is environmentally cued.

2. Recurrent Behavioral Modes of Insects

   a. Foraging, food gathering (including rituals or dances)
   b. Defense
   c. Mating (often specialized within a few individuals)
   d. Building
   e. 'Resting' (including repair, anabolism, digesting)

How many distinct metabolic states underlie these behavioral modes is not clear. Foraging, defense and building all involve energy expenditure through motility, and may represent only one metabolic state. Resting and mating clearly differ from each other and from the aforementioned states, and 'resting' may involve at least three different metabolic states, as indicated parenthetically above. Thus, five behavioral modes may be associated with as many metabolic states, though the correspondence is not 1:1.

3. Recurrent Metabolic States for Adult Mammals

   a. Sleep-rest (basal metabolic rate)
   b. Thermogenesis
   c. Digesting
   d. Exercise (energetic skeletal motion of all kinds regardless of objective)
   e. Anabolism: repair or recuperation after disease or injury, muscle hypertropy after chronic exercise (may be more than one state)
   f. Fattening
   g. Fasting

The number of behavioral modes is obviously much higher than the limited number of bioenergetic states above. Thus there is not a 1:1 correspondence between metabolic, bioenergetic states and behavioral modes. However, the metabolic states listed above may in fact each be a constellation of simpler states. In the present analysis constellations will be emphasized.

(j) State transitions. - Specification of distinct constellations of metabolic processing - the metabolic states - raises the question about the probabilities of transition from one to another. Is the ensemble of these processes ergodic? Is the trajectory through the phase space of metabolic states that of a Markovian process, determined by the present and recent past, or are the probabilities to be described otherwise? Whatever the answers to these questions, the transition from one metabolic state to another requires signals for initiation. It is therefore necessary, when dealing with biological systems, to distinguish power fluxes from informational fluxes. It is also important to know whether or not a continuous informational flux is required to hold a metabolic system in a given state, or whether the informational flux need only be a transient, required to displace the metabolic system to a new, stable operating point that it can sustain after the cessation of the informational flux.

A 'message' within the metabolic system is defined as having been received when a substantial change in mass flow across membranes or in biochemical chains has occurred, or a substantial energy transduction appears, as is involved in the conversion of ATP to muscle contraction. Messages are carried ultimately by chemicals (substrates, drugs, hormones, ions and other

metabolites), or by the neurotransmitters, norepinephrine and acetylcholine. These chemical signals may be created and distributed locally, as in the case of neurotransmitters, prostaglandins, and kallecreins, or they may be distributed generally, as in the case of hormones, through a common, hydraulic, convection channel such as cerebral spinal fluid, sap, blood, hemolymph. The distribution takes hormones to many regions of the living systems where the signal molecule is 'recognized' at specialized target areas on cell surfaces or interiors.

(k) Chemical signals. - The number of metabolites that may affect biochemical chains or the genome is unknown, but in mammalian cells the number is probably not less than 100. These substances act as intracellular communication agents. The number of systemic hormones that carry messages between cells is about forty for mammals, nine for insects and six for plants. The number of classes of local hormonal signals is not known, but is apparently small, and includes histamine, kallicreins and prostaglandins.

To the above list of chemical signals might be added the pheromones that are chemicals transmitted from one living individual to another through the air or water medium between, which, when detected, affect the behavior and metabolic state of the recipient. The number of pheromones is not known even approximately, but for any given species it is apparently small (only one).

In summary, we recognize perhaps 200 chemical signals that may command changes in energy transduction or mass flow in a mammalian system, and these represent about an equal number of intracellular metabolites and extracellular hormonal messengers. In the present analysis, only the hormonal messengers are considered, because the level of function of interest involves that of multiple organs connected by communication links involving only nerves and hormones. At the multi-organ level of function, the processes underlying bioenergetics involve very few signals. The two neurotransmitters account for the influence of nerves on organs and their secretory supply. The number of hormones affecting bioenergetics in the mammal is only six: insulin, glucagon, growth hormone, epinephrine, thyroid hormone (actually two) and glucocorticoids (actually two). Adjustments of water and mineral balance involve four additional hormones (antidiuretic hormone, thyrocalcitonin, parathyroid hormone and aldosterone). The minerals whose levels are controlled by these agents are sodium ion, potassium ion, calcium ion, and phosphate ion.

(l) Convergence of chemical signals. - The relationship between a communication flux received and a power flux achieved by a recipient cell defines metabolic command 'languages'. The large number of possible chemical signals could be taken to imply that for each there is a specific receptor (as is probably true), and for each receptor there is a distinct function. The latter assumption is probably not true, for the many chemical symbols apparently converge at the cellular level to only a few intracellular languages. A translation occurs, as has been discovered in the special case of cyclic AMP.

It is now recognized that cyclic AMP is an intermediate in the action of a large number of hormones. Examples include the activation of phosphorylase in liver and muscle by epinephrine, or in liver by glucagon; stimulation of corticosteroidogenesis by ACTH; release of some pituitary hormones by hypothalamic releasing factors; adjustment of the permeability of renal tubular membranes by antidiuretic hormone; action of insulin on glucose or amino acid uptake in muscle; action of insulin, antidiuretic hormone, ACTH, growth hormone or epinephrine on the hormonally sensitive lipase of the fat cell, etc. Apparently, in the course of evolution metabolic machinery in the cell interior became responsive at numerous points to cyclic AMP. The specificity of action of the hormones that course between cells appears to occur at cell surfaces, where the arrival of a hormone molecule is translated through the cell membrane into a flow of cyclic AMP at some region of the cell interior. Cyclic AMP has thus been called "a deputy hormone".

The multiplicity of actions of cyclic AMP within cells does not fully describe the importance of this compound. The same substance has been found to be the organizing signal for aggregations of amoebas in the slime mold, and in that case it is not merely an intracellular agent, but instead it is released by pacemaker cells and diffuses through a population of cells who orient themselves in a diffusion gradient of this substance. Differentiation occurs. In this role, cyclic AMP appears as a hormone itself of the pheromone type. In all cases cyclic AMP provides commands 'understood' by an astonishingly large variety of cells and processes.

(m) Summary of Section 1. - The essential features of metabolic systems underlying bioenergetics that appear from this analysis are relatively few. They are summarized below:

The energy-producing systems all involve multi-enzyme chains numbering from ten to 100 steps. These chains include glycolysis, the tricarboxylic acid-electron transport system and photophosphorylation.

Metabolic chains are regulated by adjustments of the amount or activity of enzymes at rate-limiting or unidirectional steps.

Control of enzymes is achieved by adjusting the rates of their synthesis through the nucleic acid-protein synthesis machinery, or through conformational activation of the enzyme by hormones or metabolites.

Metabolic chains are tightly coupled to cofactor reactions that donate or receive protons and electrons or phosphate groups.

Metabolic chains are weakly coupled to each other through shared reactants in metabolic pools.

Metabolic chains may have feedback or feedforward loops.

The flow of mass through a metabolic channel can be adjusted by changes in enzyme performance at controllable steps within the chain, or by gating the presentation of substrates to the first step in the chain through selective, membrane transports.

Hormones carry messages between cells, and affect both transports and chains.

The number of hormones involved in bioenergetic processes is small, and their messages are translated into an even smaller number of intracellular languages, of which cyclic AMP flow is the archetype.

In plants, metabolic states and behavioral modes are synonomous; in insects, the number of metabolic states and behavioral modes are the same, but there is not a 1:1 correspondence; in animals the number of behavioral modes is much larger than the number of bioenergetic, metabolic states.

2. Life Without Nervous Systems

The preceding section dealt with general aspects of cellular metabolism. As a foundation for the interpretation of animal (human) behavior we now review more particularly several forms of life in which repetitive metabolic performance is achieved without the intervention of a nervous system. These forms reveal varieties of periodic, metabolic performance achieved by the flow of chemical signals.

(a) Microorganisms and cell cycle biochemistry. - The simplest form of life known is the pleuropneumonia-like organism (PPLO). Some of these organisms are only 0.1 micra in diameter (only 10% of the diameter of the average bacterium). The mass range between these organisms and a large protozoan such as an amoeba is 10^9. In fact, a single atom is as close to PPLO as PPLO is to an amoeba in size. PPLO meets all the conditions of life: it grows and reproduces in a medium free of other cells. The metabolic machinery consists of a subset of the usual animal cell parts: a flexible membrane 100 Å thick, containing cholesterol esters; enzymes (about forty); DNA (molecular weight 45×10^6. Since it takes a strip of DNA of molecular weight about 1×10^6 to regulate the synthesis of one enzyme, it appears that almost all the available DNA in the cell is involved in synthesis of its enzymes.) PPLO organisms derive their free energy from glycolysis by the usual path, yet they contain altogether only twelve hundred large molecules. The metabolic states open to PPLO appear to be two: rest, and growth with replication. Unfortunately, the cueing for the transition, and the nature of the steady-state available to this organism (if any) is unknown. Therefore, to examine cell cycle biochemistry it has been necessary to study larger single-cell organisms such as bacteria, or animal cells separated and placed in culture. The following conclusions have emerged from such studies.

(b) Cell-cycle biochemistry. - The replication of an animal cell in the controlled and constant environment of cell culture is a cyclic process. The cycling is programmed from the cell's own genetic system, but the cycling can be controlled by environmental changes if these are present. The result is that in their normal setting, some cells replicate every ten to twelve hours, and some, never.

The cycle begins with a quiescent interphase. This phase is followed by DNA synthesis and histone synthesis, then by RNA and protein synthesis, then by mitosis, and a return to interphase. Populations of similar cells in culture do not, in general, carry out these processes synchronously. The same cells in their natural state (not in culture) are usually considered to be more differentiated, and not as likely to show the cycle, presumably because their environment provides inhibiting signals. How interphase is terminated in a constant environment, and replication initiated, is unknown. The cellular machinery appears to be inherently unstable in the interphase, and the length of interphase is exceedingly variable even for cells of the same type in a constant environment. It is hypothesized that cells must accumulate an 'initiator' that is produced steadily, until a threshold is reached, and DNA replication is triggered. DNA replication is an energy-consuming process. Thus the cell cycle is postulated to be a relaxation oscillation with escapement. After cell division the daughter cells begin to reaccumulate the DNA replication initiator. It is important to note that under the conditions of constant environment, the only stable state available to cells in culture is that of cyclical replication of the relaxation oscillator (limit cycle) type.

The demonstration that the basic cell replication cycle has the characteristics of a limit cycle oscillation when the cell is in a constant environment supports the provisional generalization that metabolic systems may be only marginally stable, or stable only in the limit cycle condition of Poincare, because nonconservative, nonlinear process constellations are involved. The internal inhibitions that come from repressors of genetic activity are necessary for the limit-cycle stability of the cell, as may be judged from the catastrophic destruction of cells that occurs after virus infection, which results in the introduction of unrestrained nucleic acid and protein synthesis. The virus is replicated, but the cell dies. From the frame of reference of the virus particle a different cycle of infection, replication, cell rupture, disgorgement of amplified members of viruses, reinfection of new cells, now emerges. Here again the overall process has limit-cycle characteristics, but with respect to the process of virus replication in this case, not with respect to cell replication.

The possible general conclusion from these primitive systems in culture is that many biological processes at the primitive level may have oscillatory, limit-cycle characteristics, and when two processes are mutually exclusive (such as virus replication and cell replication) the process with the lesser inhibition will dominate. Cell replication is repressed, virus replication is uncontrolled. Thus, viruses invade and destroy cells and their life cycle supersedes that of the cell cycle biochemistry.

(c) Metabolic steady-states in non-replicating, normal cells. - B. Chance has demonstrated that normal cells have at least some oscillatory, steady-state processes. Oscillatory behavior has also been noted in single neurons, and other single-cell organisms have been shown to have circadian rhythms in luminescence or photosynthetic capacity, etc. Goodwin has sought for the bulk thermodynamic properties of biochemical oscillators, and has proposed a theoretical account of these. The inhibitory effects of gene repression are viewed as an example of strong coupling between oscillatory systems, whereas

weak coupling occurs when systems share a common pool of metabolites, as is extremely common. Whether or not cells have a number of possible oscillatory steady states, or only a few is still an open question, but in any case the transition from one such state to another can be brought about by environmental changes in nutrients, or by other chemical agents, as previously described. According to Goodwin, either entrainment or subharmonic oscillation occurs to shift the ensemble of nonlinear oscillators to new steady-state frequencies. He raises the question of whether or not cells shifted to new steady states are stable in those states after cessation of the environmental stimulus. As will be mentioned below, in plant cells the answer appears to be negative.

(d) Metabolic systems in plants. - Plant physiology provides sharp insight into metabolic systems and their interaction with the environment, because plants are relatively simple (few organs, no nervous systems, few hormones) and because changes in state are clearly cued by the environment, as can be seen in seasonal growth, flowering and abscission with dormancy. Before considering seasonal periodicity in the metabolic performance of plants, however, it is important to detail the circadian rhythms observed in them, many of which are endogenous, and are therefore not cued by the environment, but only entrained by geophysical variables with 24-hour rhythms.

(e) Endogenous circadian rhythms in plants. - In 1729 de Mairan discovered that rhythmic leaf movement in plants persisted in a darkened room for many days before subsiding. Since then many endogenous rhythms in plants have been discovered with periods close to, but not exactly equal to 24 hours. The free-running periods range from 21 to 28 hours. All of them are entrained to precise 24-hour rhythms under natural conditions, and almost all of them show damping under constant environmental conditions. That is, the rhythm subsides when the plant is put into a constant environment. Not all plants studied have circadian rhythms. Sometimes the rhythm will be discovered in one of a pair of closely related species, but will not be present in the other. In the Bryophyta group, no circadian rhythms at all have been discovered. Therefore, rhythms with this period are not universal, though they are very common.

Circadian rhythms in plants may be entrained by geophysical variations in light, temperature, or the partial pressure of oxygen. The effects may be a phase shift or a change in steady-state period. In general, an entrained oscillatory system returns to its initial value immediately after the organism is returned to a uniform environment (contrary to the expectations of Goodwin). However, in one case an entrained rhythm was found to persist for several days after return of the organism (Hydrodictyon) to a uniform environment.

Entrainment often manifests itself through frequency demultiplication so that a subharmonic oscillation occurs.

All circadian rhythms, free-running or entrained, require energy inputs since all natural systems are lossy. In the absence of oxygen, or at low temperatures, plant rhythms are abolished. Even bright light may abolish the oscillations, though the mechanism is not known. Apparently bright light

drives the rhythms to some specific phase in the cycles, and holds them there.

When a species reveals multiple endogenous circadian rhythms, their phase relations may bear on the health of the organism. The phase relations of free-running cycles are altered by entrainment from an environmental signal. Even brief environmental changes affect the phase relationships. The phase effects have suggested that the cycles are relaxation oscillators, with pulsed energy inputs provided through escapements, and an energy independent recovery in another portion of the cycle. Though it is not inconsistent with this view, it is interesting that the parts of the cycles over which low temperature or anaerobic conditions can be applied without delaying the phase are quite different, although both types of treatment might be expected to have common effects in reducing energy inputs. The pathways by which they do so are different.

The periods of circadian rhythms in plants and animals may be readily entrained by light, but they are surprisingly temperature compensated, though not perfectly so.

In animals, the circadian rhythms may arise from a few specialized cells, as is clearly the case in the cockroach, but in at least some multi-cellular plants each cell has its own rhythm. The location of the rhythm within the cell has not been achieved. Enucleated cells still have the rhythm in their cytoplasms, but if the nucleus of a cell entrained to one rhythm is transplanted to an enucleated cell entrained to a different rhythm, the nuclear rhythm is imposed when the hybrid cell is placed in uniform conditions. The basic oscillatory system is apparently not in the nucleus, but some entraining influence can be transmitted via the nucleus. Nevertheless, phase shifting by light can be achieved whether or not the nucleus is present, so the nucleus is not obligatory in the coupling to the environmental stimulus.

At the present time it appears that many, but not all cells, have an oscillatory chemical system with a period between 21 and 28 hours, located in the cytoplasm or in the organelles, but not in the nucleus. This oscillation further supports the notion that metabolic steady states are often oscillatory and not static at the lower levels of organization. The nature of periodic changes at higher levels of organization is considered below.

(f) Seasonal periodicity in metabolic performance of plants. - In the round of a year the deciduous plants show seasonal growth, flowering and abscission. These changes in metabolic state are cued from the environment, and involve several hormones within the plant. The hormones known to be involved are: giberellins, auxins, florigen, cytokinins, abscissic acid and ethylene. These hormones account for geotropism, phototropism, growth, flowering, and the dropping of leaves. The strongest environment triggering signal for these processes are day length, and this signal is translated into hormonal release by means of the pigment phytochrome. Some species of plant require chilling to overcome seed dormancy and to achieve normal growth rates. Therefore, changes in temperature are also triggering signals. It is remarkable that some plants will show slowed growth for as long as ten years if they are kept in a warm greenhouse, but when they are pulsed with cold, normal growth is resumed.

Plant hormones are most effective in the regions where they are produced, but they do travel to remote parts of the plants as well, probably via the plasmodesmata. Plasmodesmata are cytoplasmic bridges between cells that allow for passage of substances and electric currents from one cell to another at their junctions.

The triggering of great shifts in the metabolism of plant organs by changes in temperature or the length of the day as detected by a pigment, and transduced into hormonal flow within the plant, favors an epochal view of metabolism, with hormones acting as switching signals. In this view, the plant moves from one state to another, but it is not clear that any of the states is a steady state. Annual plants go through one trajectory from seed germination to seed production and death in one year. The life cycle then is a succession of individual forms, none of which has any steady state available to it on the time scale of a year. On the time scale of a day, however, the metabolic system even of the annual plants may be considered stable. These facts indicate the usefulness of attaching to each biological process of interest a specification of the time domain appropriate to it, and to indicate this domain as a period in the case of repetitive phenomena.

A striking feature of the coupling between plant metabolic processes and the environment is that the environmental changes often need be present only a short time, to act as trigger, to commit the metabolic system to a substantial, often relatively persistent shift in state. Thus it may be concluded provisionally that plants have quasi-stable metabolic states in which they may remain with marginal stability, until the next environmental trigger disturbs the stability, causes a hormonal flow that displaces the operating points of the metabolic chains within the plant organs in such a way that a new regime of marginal stability emerges. The number of possible marginally stable states for plants is not high, as indicated previously. Although hormones appear to act to express the triggering signal as a change in operating point of the metabolic systems, so the transition to a new state occurs, some of the hormones such as cytokinins, are clearly not triggers, and these appear to be needed continuously for bud formation. The cytokinins are part of transfer RNA, even in calf liver, where they also occur. It is not certain that their action as hormones is related to their appearance in soluble RNA.

From the above discussion it will be noted that there appears to be a discrepancy between the transition of a plant from one marginally stable metabolic state to another on the seasonal time scale, compared to the transition from one circadian rhythmic state to another by entrainment from the geophysical environment of the plant. In the former case, environmental signals may act as triggers, whereas in the latter case they must be present continuously to hold the plant in the new state. It is as if in the circadian domain only one marginally stable state exists, whereas in longer time domains several marginally stable states exist. However, these conclusions cannot be made with certainty on the basis of the present evidence.

(g) Hormones and time domains. - The hormones of complex plants act in more than one time domain, and have more than one effect. For example, auxins, giberellins, and ethylene all affect protein synthesis. Auxins increase ribo-

somal RNA after a lag, but may stimulate growth without a lag, thus operating in two time domains. Giberellins cause an increase in de novo synthesis of α amylase, through an increase in a specific mRNA. The hormone acts directly on nucleic (even on isolated nuclei) and a lag is involved. Only one time domain is involved, perhaps, but the effects are multiple. Flowering starts, and peroxidase is inhibited.

The peroxidases of plants inhibit plant growth. Since auxins, cytokinins and ethylene also affect the levels of peroxidase, as does light or wounding, it can be seen that extensive convergence occurs at metabolic control points in plants, of which peroxidase is one. No fewer than four classes of chemicals and one or more physical stimuli affect the same enzymatic activity. It is not known whether a cyclic AMP is involved in the convergence of signals within plant cells, but in any case, whatever the intracellular signal traffic, convergence is seen in plant cells as in animal cells. That is, at the cell boundary a translation of environmental signals occurs so that the messages carried into the cell interior are conveyed by relatively few languages.

(h) General principles of frequency demultiplication and frequency multiplication in nonlinear systems. - The fact that hormones have effects on processes that occur in multiple time domains does not mean that the periods of these processes will appear in the pattern of hormone levels. Ensembles of nonlinear oscillators can be coupled to give frequency demultiplication (a good example is heart block, in which case the atrial-ventricular node responds to every second or third cycle of the sino-atrial node) or frequency multiplication (as occurs in any unsymmetric detector network).
Thus, it is impossible in ensembles of nonlinear oscillators to establish chains of causality solely by matching frequencies. A spectroscopic approach to such ensembles can only define the time domain in which large scale power fluxes occur, but it cannot establish which communication fluxes were responsible for initiating the power fluxes. The possibility of frequency demultiplication or frequency multiplication rules out the unique assignment of causality through frequency matching. Nevertheless, a spectroscopic approach to biooscillators helps to describe plant processes and to determine whether or not the regimes of stability of chemical oscillators consist of several discrete modes, or instead, of many constellations of periods of the separate oscillators.

(i) Summary of Section 2. - From the foregoing discussion several conclusions emerge:

Periodic performance of metabolic systems appears at the lowest level of organization that meets the criteria of life. At that level the cycle is the cycle of cell replication in constant environment.

Periodic function of larger scale metabolic systems can be seen in the circadian and seasonal rhythms of plants. These periodic functions, as in the case of the lower organisms, occur without the intervention of nervous systems. Chemical signals are involved, and in the case of plants hormones have evolved to achieve communication among cells and organs.

Circadian rhythms in plants are highly damped, are easily entrained by environmental stimuli, and do not persist at the entrained rhythms once the environmental oscillator is removed.

Seasonal cycles in plants may require pulsed, triggering signals from the environment that recur periodically over a year, but they do not require continuously acting stimuli from the environment.

Plants have about six classes of hormones, that achieve regulation of growth, flowering, leaf-dropping and dormancy, as well as geotropism and phototropism. Some of these hormones are continuously active, and some appear to act as triggering signals.

The study of plant hormones reveals that a single hormone can affect processes occurring in multiple time domains.

Because ensembles of coupled, nonlinear oscillators show both frequency demultiplication and frequency multiplication, as well as the possible adoption of an intermediate frequency shared by no individual oscillator, it requires great caution to use a spectroscopic approach to relate oscillating informational fluxes to oscillating power fluxes in living systems.

3. Time Domain of Cellular Processes

Chemical processing by cells is very rapid. Individual reactions at single steps in metabolic chains occur in the time domain of 10^{-9} seconds. At the level of individual cells such as E. coli growing in an optimum, uniform medium with a cell cycle period of 20 minutes, the chain process rate for the various important molecular species have been found to be as shown in Table 3.

Table 3

Chemical Component	No. of Molecules per E. Coli Cell	No. of Molecules Synthesized per Second	% of Total Biosynthetic Energy Required
DNA	4	0.0033	2.5
RNA	15,000	13	3.1
Protein	1,700,000	1,400	88.0
Lipids	15,000,000	12,500	3.7
Polysaccharides	39,000	33	2.7

It can be seen that the cellular chemical machinery is capable of almost fantastic rates of synthesis. In the last stage of construction of each protein at least 100 covalent peptide bonds must be formed at the ribosome. Since 1,400 protein molecules are synthesized per second, the ribosomes alone must be carrying out peptide bond formation at the rate of 140,000 reactions per second, minimally. The coordination of activity of metabolic chains

therefore occurs in a time domain ranging from nanoseconds to no more than minutes during the cell replication cycle. Of course, in mammalian tissues and organs, it is possible that specialized cell activities require the operation and adjustment of biochemical chains in longer time domains. In fact, however, those chain adjustments that can be made without requiring de novo synthesis of enzymes seem always to occur in a matter of seconds or minutes.

Adjustments of the actions of metabolic chains through enzyme synthesis in mammalian cells is much slower than the process described above. It appears that de novo synthesis of an enzyme requires a time lag of about one to four hours from receipt of the triggering signal to first appearance of increased rates of protein synthesis. For example, induction of synthesis of alanine amino transferase by cortisol in liver leads to a first detectable increase in enzymatic activity four hours after the treatment with the hormone. Similarly, increases in oxygen consumption provoked by thyroid hormones first appear only hours or days after initiation of the hormonal signal (an increase in levels of free thyroxin), and the effects persist for similar times after withdrawal of the signal. The longest lags or delays in chemical processes affected by hormones therefore appear to be in the domain of hours to one day.

Many chemical processes that occur over a period of hours appear to be 'locked on' by a hormonal transient, which, after a lag, starts up a plant process that may then relax only slowly over a period of many hours, after the triggering hormonal signal is released. In other cases, the process is more tightly coupled to the immediate level of the hormone, and the relaxation of a hormonally-stimulated process follows the relaxation of the concentration of the hormone itself. In such cases, the hormone is needed continuously for the process to occur. An example of such a case is the stimulation of cortisol secretion by the adrenal with corticotropin.

(a) Permissive action of hormones. - The specification of the time domain in which a hormone may act is made extremely difficult by the 'permissive' nature of the action of some hormones. For example, thyroid hormones affect the sensitivity of many processes to catecholamines. Similarly, cortisol affects the responsiveness of the hormonally-sensitive lipase of fat cells to the action of other agents such as epinephrine or growth hormone. In these cases a sustained presence of one hormone is necessary to set the parameters of the kinetic systems involved so as to make transient responses to other hormones possible. The hormones required in a sustained way may have no direct effect alone on the processes. Thus, the time domain in which they act on those processes is essentially infinite, and the process change occurs only in the time domain characteristic of the second hormonal system that produces the transients. Hormones that have permissive effects with respect to some processes, may provide the transient signals to other processes when their levels change.

It is not yet clear whether or not the permissive action of hormones is a dose-dependent action, or merely a threshold effect. Tentatively it appears to be the latter. This arrangement means that a given hormone such as cortisol can support the lipolytic action of epinephrine in a steady way while changes

in the plasma concentration of cortisol may be the proximate cause of changes in other processes.

Because of the multiplicity of loose couplings among metabolic chains, many indirect effects of hormonal action are possible and are observed. It is believed by many, for example, that an increase in the level of free fatty acids in blood decreases hepatic glycolysis and increases hepatic gluconeogenesis. If indeed, free fatty acids do control the operation of these two great metabolic chains in the liver, then an indirect effect on those chains may be expected from the actions of insulin, epinephrine, the sympathetic nervous system, growth hormone, thyroid hormone, glucocorticoids and even antidiuretic hormone, all of which affect the rate of secretion of free fatty acids into blood from fat. The time domains in which the effects lie will depend both upon the primary hormonal action, and the secondary action of free fatty acids on the hepatic processes. In such networks of coupled metabolic actions, a spectroscopic approach offers little hope of establishing causalities, as previously pointed out. What the spectroscopic approach can do is to identify periodic power fluxes, if they are present, and therefore to launch the quest for the informational flux initiating or sustaining the power flux. The information flux need not be periodic, even if the power flux is stabilized in a limit cycle mode.

4. The Structural Level Defining a Metabolic State

Up to this point this discussion has emphasized the metabolic chains and membrane transport systems that underlie bioenergetics. However, cellular systems of chains and transports do not represent the lowest level of structure at which the metabolic state is defined. The reason they do not is that substrate supplied to cells, or the delivery of hormones to them, as well as the removal of their secretory or metabolic products requires the presence of a convection system close to the cell. Therefore, the minimal structural unit determining a metabolic state is a cluster of cells with the circulatory (e.g., capillary) and lymphatic hydraulic connections to them and the regional nerve supply. In the case of cardiac muscle there is a capillary for each muscle cell. In the case of skeletal muscle, the number of open capillaries is variable. In the extreme the ratio of capillaries to active cells may be as large as 2:1. Generally, however, the ratio is smaller than this. Alterations in that ratio which are achieved by adjustments of the circulation clearly have profound effects on metabolic systems through delivery of substrates or removal of products, as well as through the conveyance of signal molecules to and from regions of active metabolism. Both the informational fluxes and the power fluxes are dependent upon the circulation in most cases. In some cases the informational fluxes are determined by nerves, and the neurotransmitters released at the surfaces of active cells. In no instance is the informational flux through nerves thought to be the only means of influencing a metabolic system.

In the case of the liver, the minimal structural unit would be the lobule. The functional unit of the liver, portrayed earlier, has all the control features found in the whole liver, and any unit less than this lacks one or an-

other of the important influences to which the liver may be subjected. In the case of the kidney, the minimal metabolic unit is the whole nephron.

It can be seen that a full description of the metabolic state of an animal requires a description of regional blood flows, nerve traffic, hormonal supply and substrate supply in arterial blood. A region of active metabolism whose circulation is interrupted not only ceases to carry out the active metabolism, but also ceases to have an influence on the general metabolic state of the animal. Thus the coupling between metabolic chains, cell membrane transports and the microcirculation is very tight. The tightness of this coupling can be noted in a highly flow regulated organ such as the brain.

The brain of man consumes about 20 watts of power (0.3 kilocalories/minute) which represents about 25% of the basal metabolic rate. This power requirement is met by a blood flow of about 40-50 ml/100 gm brain/minute. The extraction of oxygen from the arterial blood is approximately 1 ml of oxygen removed from each 14 ml of blood delivered. The oxygen consumption of the brain is quite steady at 3 ml O_2/100 gm brain/minute. The substrate for power production in the brain is exclusively glucose under normal conditions. Ninety-five per cent of the glucose is oxidized to carbon dioxide and water, but five per cent produces lactate, a result that suggests that the brain is always slightly hypoxic. The lactate being produced acidifies the cerebral blood. Regulation of the resistance vessels in the brain is accomplished by carbon dioxide, which is the product of aerobic metabolism of glucose, and by lactate (hydrogen ion concentration) which is the product of anaerobic metabolism of glucose. Increased CO_2 or H^+ both are powerful dilators of cerebral resistance vessels. The sequence of metabolic events in the brain is given in one description as follows:

If blood flow to a region of the brain is slowed or stops, the neurons and the glial cells rapidly use up the available oxygen.[1] The metabolism of glucose shifts within five seconds to the anaerobic mode and lactate production increases. Hydrogen ion concentration increases and the sphincters dilate so that blood flow is resumed. The capillaries of the cerebral circulation appear to have a periodic function with an opening and closing cycle of 20 seconds. Autoregulation of cerebral blood flow is prominent. The flow rate is independent of the pressure drop across the brain. If blood pressure is raised and flow begins to increase, tissue oxygen tension increases, and the anaerobic metabolism of glucose is decreased and lactate production drops. The cerebral vessels then constrict, resistance increases and flow is reduced back to normal in the face of an increased pressure difference. The autoregulation of cerebral blood flow depends upon the Crabtree effect (decreased oxygen supply causes increased anaerobic metabolism of glucose) and the Pasteur effect (increased oxygen supply inhibits glycolysis). These two effects are mediated by the levels of ADP and ATP within neurons. In the absence of oxygen, ATP levels fall, ADP levels rise, respiration is inhibited and gly-

[1] The total role of the glial cells in mediating diffusion and transport from capillaries to neurons in the CNS is not adequately defined. Illustrating, glia is sensitive to CO_2, possibly as a response to pH changes.

colysis is facilitated. In no other tissue is the coupling between metabolic performance and the microcirculation as tight as it is in the brain. The coupling is achieved without the intervention of the autonomic nervous system, but depends entirely upon metabolites.

5. Bioenergetics in Animals at the Organ System Level

Bioenergetics in animals involves fourteen major processes in three major organs. The organs are liver (and to a much lesser extent, kidney), muscle and fat. The fourteen processes are listed below in Table 4.

Table 4

Bioenergetic Processes

Circulation-coupled transport processes across cell membranes

1. Glucose uptake
2. Fatty acid uptake
3. Amino acid uptakes
4. Glucose release
5. Fatty acid release
6. Amino acid release

Metabolic chain processes within cells

7. Glycogen synthesis
8. Glycogenolysis
9. Fat synthesis
10. Lipolysis
11. Protein breakdown (protein synthesis goes on independently of bioenergetic requirements of the whole animal and is not listed here as a separate process)
12. Gluconeogenesis (glucose synthesis from certain amino acids, lactate, citrate or glycerol)
13. Oxidative phosphorylation by mitochondria
14. Glucose utilization via glycolysis

(a) Substrate loading of metabolic systems. - All of the fourteen processes listed in Table 4 can be driven on a thermodynamic basis by substrate loading (the law of mass action), until a rate-limiting step is saturated and the process velocity becomes maximum and constant. Such performance may be likened to that of a freeway system without traffic lights: the traffic ebbs and flows at different times of the day depending on the rate of presentation of automobiles in the on and off ramps of the freeway system. A substantial amount of metabolic following action occurs in this manner. Nevertheless, a traffic light, lane control involving neural and hormonal signals has evolved that optimizes the performance of the system beyond what could be achieved by

mass action effects alone. For the processes of bioenergetics, in mammals, six hormones appear to be involved. One of these is a hormone of 'feasting' and adjusts the metabolic system when substrate supply through the gastrointestinal tract is plentiful. That hormone is insulin. Five of the hormones are hormones of 'famine'. They adjust the metabolic system during periods of acute or prolonged fasting. All six of the hormones have slow actions (in the time domain of hours or days) exerted through effects on enzyme synthesis. Four of the six have additional, important actions in the shorter time domain of seconds to minutes (insulin, glucagon, epinephrine and growth hormone). Table 5, below, lists the metabolic processes in muscle (M), fat (F) or liver (L) affected in the short time domain by the four hormones just referred to, and in the long time domain by glucocorticoids and thyroid hormones. The long time domain effects of the four hormones with faster actions as well are not listed. Each process is preceded by a capital letter designating the organ in which the effect occurs, and a number designating the process shown in Table 4 that is prominently affected.

Table 5

Bioenergetic Processes Affected by Hormones

Hormone	Cellular Process
Insulin	M1, F1, M3
Glucagon	M11, H12, H8, H4, F10
Epinephrine	H8, M8, F10
Growth Hormone	M3, F10 (and protein synthesis is stimulated)
Glucocorticoids	F10, M11, H12, and inhibits M1, F1, and stabilizes lysosomal membranes
Thyroid Hormones	M13 and increases sensitivity to neurotransmitters and epinephrine

The effects of the various hormones on bioenergetic processes can be translated into words in the form of the following commands:

Insulin: "The supply of glucose is plentiful, take it up, burn it, and store the excess as fat."

Glucagon: "The supply of glucose is running low; the liver should add some from the glycogen stores to the blood."

Growth Hormone: "The supply of glucose is running low; the fat cells should add fatty acids to blood for fuel."

Epinephrine: "The supply of bioenergetic substrates is still running too low; do what growth hormone and glucagon have commanded, but do it more and faster."

Cortisol: "We have a problem here; in spite of the best efforts to date the bioenergetic substrates are still too scarce; tear down the muscles and use the amino acids of the proteins as fuel to make blood sugar for the brain."

The somewhat whimsical presentation above captures the sense of the metabolic commands given through the hormonal system to the three chief metabolic organs during feasting, or during a progressively severe famine. This listing of commands and responsive processes completes the description of the metabolic machinery and its control arrangements. The final section will be devoted to some general comments about the overall performance of the bioenergetic systems in animals.

6. Concluding Comments About Bioenergetic Systems of Animals

From a physical point of view, certain questions arise concerning the operation of the hierarchical bioenergetic mechanisms described above in this report. Those questions concern the degree of stability of the bioenergetic system and the form of that stability, the number of 'languages' used for metabolic commands, and the relationship between intake regulation and metabolic follower action. Each of the questions will be considered here in turn.

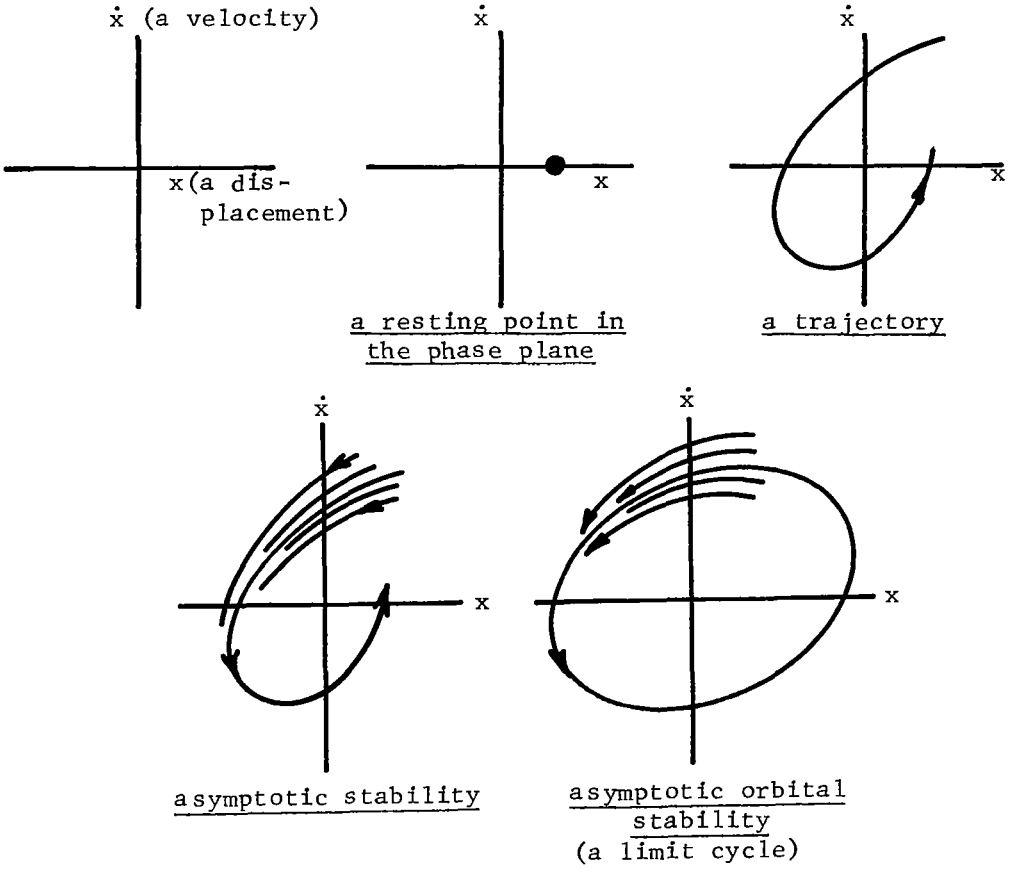
(a) Stability. - As described above, the bioenergetic systems involve biochemical chains, membrane transports and the microcirculation. Signals are transmitted in the form of metabolites, hormones or nerve impulses and transmitters. The coupling between the microcirculation and the metabolizing cells is close. The actions of chemical signals on the metabolic systems involve lags and delays, feedback and feedforward, numerous parametric adjustments, and a large number of reactions with highly nonlinear kinetics. It is obvious that the bioenergetic systems of animals are open nonconservative systems. Furthermore, it has been clearly established that although many of the reactions in the metabolic chains operate near equilibrium, key, controlling processes (including structures) and reactions operate at points quite far removed from equilibrium. To all of these properties can be added the obvious one that metabolic systems are in fact stable. Organisms may live as long as a century or more.

Some generalizations about the stability of nonlinear systems can be made. Structural stability is the property of a physical system, linear or nonlinear, such that the qualitative nature of its operation remains unchanged if parameters of the system are subject to small variations. Many aspects of biological systems clearly lack structural stability. The nerve impulse is a particularly well known example, as is the autonomous pacemaker of the heart. In both these cases a very slight change in the conductance parameters of the cell membranes (which are voltage-dependent terms) leads to explosive changes in the physical properties of the membrane. The performance is that of a relaxation oscillator with an escapement, and structural stability is not present. It is more difficult to judge whether or not metabolic systems have structural stability. With respect to some parameters it seems likely that structural stability is present. For example, in a model of corticosteroidogenesis, it was demonstrated that some of the parameters can be varied plus or minus 100% around their nominally normal values without affecting the stability of the overall kinetic network. In contrast, a slight variation in other parameters leads to the displacement of the whole network to an extreme position of maximal or minimal function. Overall then it is

unlikely that metabolic systems have structural stability. Therefore we must seek elsewhere for the origin of their dynamic stability.

Since nonlinear systems may have multiple equilibrium[1] points available to them, the question arises as to whether or not they can be displaced from one equilibrium operating point to another. Since metabolic systems are sustained but dissipative open systems, resting equilibria are not reached during life. Instead, stability is characterized as a nonlinear dynamic (e.g., fluctuating) steady state.

From the work of Poincare and Liapunov, it is known that the stability regime of such a system is not with regard to a point in a phase plane of displacement versus velocity, but rather with regard to the nature of the possible trajectories of the system in that phase plane. First, a trajectory can decay asymptotically to a point. This constitutes resting equilibrium, which we have indicated does not represent the living state. Second, if a trajectory is not closed in the phase plane, and it and all its neighboring trajectories converge to a limiting trajectory, the motion is said to be asymptotically stable. Further if the trajectory is closed and representative of asymptotic stability, then the closed trajectory is known as a limit cycle.[2]



[1]More properly these should be regarded as singular points of the motion.

[2]A limit cycle is essentially asymptotically orbitally stable. Orbital stability has to do with neighboring trajectories being non-divergent.

Can we find evidence of limit cycles in living systems? The answer is clearly yes, and examples have already been presented from cell cycle biochemistry. In the adult mammal, we do not yet know with certainty whether the chief metabolic variables in blood (glucose, amino acids, fatty acids, lactate) show persistent oscillatory performance. However, as described for the brain cell-capillary metabolic unit, oscillatory behavior has been reported in blood variables. Much more experimental work is needed to determine whether or not limit cycle oscillations truly characterize the stability regime of metabolic systems generally, but the conclusion seems almost inescapable that such a characterization will emerge.

Assuming for the moment that metabolic systems have limit cycle stability, the question then arises as to whether the whole animal has a stability regime characterized by a single limit cycle, or by a constellation of separate limit cycles for each process. According to the analysis presented here, the number of distinctly different metabolic states that can be discerned at high levels of organization appears to be few. In the case of the PPLO only two such states have been discovered (rest and grow-divide); in the case of the higher plants, insects and animals the number of states appears to lie somewhere between six and twelve. We are tempted to associate these states with singular states of the system. Under these circumstances, it appears that the stability regime for metabolic systems consists of a few constellations of stable orbits of coupled, nonlinear, limit-cycle process oscillators. In this view, hormones may be described as communications signals that transduce internal and external environmental influences into changed operating points of the involved oscillators, so that a new stable constellation is reached. What remains unclear, however, is whether or not the new stability regime can be held when the environmental stimulus is withdrawn. The available data indicate that in some instances the answer may be yes, but in most instances there is return to a simple pattern of constellations. The conclusion would then be that there are not more than a few basic, stable metabolic states, and the other apparent states are in fact transient unstable displacements from the basic ones.[1] They can be maintained only by the continuous application of a disturbance. These important points have certainly not been settled with respect to mammalian systems, although perhaps they are coming into focus for simpler plant and animal systems. Even though, as noted several times above, the spectroscopic approach to description of biological systems has great difficulty in helping to establish the causal relations between information fluxes and power fluxes such an approach can resolve the question of whether or or not there is more than one constellation of stable orbits in metabolic systems. The biospectroscopic approach to bioenergetics may be expected to open the way to a new understanding of these systems and their trajectories that define the life of a plant or animal.

[1]For example, some basic states may be eat from surround-grow-divide, photosynthesize, eat from stores, rest (sleep)-wake, rest-forrage. The transient states may be eat (in higher mammals), aggregate, migrate, hibernate, sporulate.

(b) Metabolic languages. - Adjustments in the metabolic activity of a region can be achieved by agents that act directly upon or within cells, or upon the microcirculation. The agents are of five types, and each type may be considered to constitute a chemical language interpretable in metabolic systems. They are listed below:

Intracellular signals
- Cyclic AMP

Metabolites
- Carbon dioxide
- Lactic acid

Tissue substances released locally
- Histamine
- Bradykinin
- Kallidin
- 5-Hydroxytryptamine
- Renin
- Prostaglandins

Neural transmitters
- Norepinephrine
- Acetylcholine

Hormones
- Insulin
- Growth hormone
- Glucagon
- Epinephrine
- Cortisol
- Thyroxin

As far as is known, the ions that have profound effects on membrane stability and bioelectric phenomena (sodium, potassium, calcium, chloride) are not carriers of information fluxes to metabolic systems, except insofar as the sodium and potassium fluxes associated with the action potential of nerves and muscles is involved in metabolism. The calcium ion concentration of blood is so severely stabilized that it is not a carrier of messages.

(c) Intake regulation and metabolic follower action. - Although metabolic chains contain feedback loops, it should be clear from the preceding discussion that the stability of metabolic systems is not that of a negative feedback system. The stability arises out of the coupling of an ensemble of nonlinear, limit-cycle oscillators, and the state variables adopt stable values, presumably in the limit cycle mode. The metabolic state variables are not apparently controlled by servosystems that contain reference inputs, comparator-like actions, error activation, and amplification. Some metabolic loops have these characteristics, but the overall system is only sparsely supplied with feedback loops. Thus changes in metabolic states are brought about by follower actions as substrate loads are placed on the system. Substrate

detectors (such as detectors of glucose and amino acids) reside in the brain and in other organs (such as the pancreas), and these detectors influence hormonal flows, which in turn change the parameters in the ensemble of oscillators. A new constellation of stable orbits of limit cycles presumably results. In such a system, stability can be achieved in the face of almost random inputs.[1] However, some intake regulation has been added, and the overall metabolic system is protected as a result.

Just as dehydration is a strong stimulus to thirst and the behavioral seeking and drinking of water, so is hypoglycemia a strong stimulus to hunger, and the behavior appropriate to finding food. In both cases the brain appears to be directly sensitive to the chemical composition of blood. The final elaboration of the stability of metabolic systems thus involves in the higher animals, a behavioral output of brain that leads to the seeking of nutrients at appropriate times. Such food seeking is itself periodic and it represents a still higher level cycle in metabolic performance.[2] The longest-term metabolic regulations involve the cooperation of intake regulation and metabolic follower action, and these concern the regulation of body weight over the period of months to years. Since the first law of thermodynamics must be satisfied for a duty cycle, the regulation of body weight involves the achievement of caloric and water balance through adjustments of intake, and activity. At these long time domains, the stability is achieved by the system behavior at the highest level of organization - that of the whole animal individual. The metabolic follower action of the internal metabolic system is helpless to offset the caloric imbalance produced by forced feeding (as in the fatting of geese in France), though it has a slightly greater range of possibilities with respect to the forced intake of water. In any case, stability of body weight ultimately involves appetites and intake regulations. The metabolic state drives the behavior, but, conversely, inappropriate behavior (gluttony) can substrate-load the metabolic system so that body weight regulation is lost and obesity, for example, results.

[1] One presumes that by selection and evolution, the dynamic systems have developed into a remarkably stable autoregulatory ensemble.

[2] One should carefully note the stability chain. The environmental boundary condition is, say, availability of food, or nonavailability of food. In the former case, a limit cycle of eat-non-eat emerged. The chain is somewhat longer. For example, in a gorilla it might be stripped to wake-roam and eat-rest-roam and eat-sleep. The act of eating itself, locally, appears to be a transient. The animal must actually be faced by food to eat, and if not there, he doesn't eat. However, having 'transiently' eaten, he doesn't eat again until some subsequent period, where again he will apparently go through a transient. On the other hand, with no food available, a different limit cycle emerges. Nominally it involves eat from stores-roam. A difference emerges, generally in the roaming range, when food is not available, and in the motor patterns of metabolic expenditure, as compared to when food is available.

Intake regulation appears in microorganisms through adjustments in cell membrane permeability according to the accumulation of metabolic products within cells. In those instances the regulation is achieved through feedback loops and complex behavior is not involved.

For man the prime mover of the metabolic system is the willing into action of skeletal muscle masses. The resulting consumption of oxygen, glucose and fatty acids commits the system to adopt new states through commands issued in the five languages described above. Those new states ultimately commit the whole animal into behavioral modes seeking rest, food and water.

Bibliography

The metabolic principles upon which this section is based can be found in the following sources. Each source is a book that extensively covers the various aspects of metabolism where documented work is available. The primary documentation of the specific metabolic studies used in this report, may be found in the bibliographies included in the references.

1. Racker, E., "Mechanisms in Bioenergetics", Advanced Biochemistry Series, Academic Press, 1965.

2. Lehninger, A. L., "Bioenergetics: The Molecular Basis of Biological Energy Transformations", W. A. Benjamin, Inc., 1965.

3. Wilkins, M. B. (Editor), "Physiology of Plant Growth and Development", McGraw-Hill, 1969.

4. Webber, G., "Advances in Enzyme Regulation", Vol. 6, Pergamon Press, 1968.

5. Bittar, E. E. and Bittar, N., "The Biological Basis of Medicine", Vol. 2, Academic Press, 1968.

6. Dickens, F., Randle, P. J. and Whelan, W. J., "Carbohydrate Metabolism and Its Disorders", Vol. 1, Academic Press, 1968.

7. Symposium on Cell Cycle Biochemistry, Federation Proceedings, Vol. 28: 1754-1789, 1969.

X. THE ROLE OF THE NERVOUS SYSTEM IN THE ORGANIZATION AND INTEGRATION OF LIVING SYSTEMS

1. Cellular Dynamics

Since the 1839 papers of Matthias Schleiden and Theodor Schwann, the cell has been regarded as the fundamental structural unit of plants and animals. During subsequent years it has been found that the cell carries on many processes, most of them serving to maintain or modify its structure (1). Among these processes are the production and storage of energy, the formation and maintenance of boundaries, the breakdown and assembly of specific molecules, the relative movement of large portions of the cell, the reproduction of the cell, the modification of the cell to carry out specialized functions with respect to other cells, and the regulation of all these cellular processes.

The mechanisms by which regulation of these cellular activities are achieved are presently being examined (2). The underlying principle is that the chemical-physical processes of the cell are in non-equilibrium steady states, and that the rates of most of these processes can be varied by specific enzymes. Regulation can be achieved by accumulation of an end product which inhibits the enzyme of the rate-limiting step in a chain of reactions (3, 4). It can also be achieved by inhibition of the synthesis of an enzyme which speeds the formation of a given product, the inhibition being worked either by that product or by another molecule (5, 6). The regulatory mechanisms are far from being understood (4). There is increasing evidence that many of the regulated processes depend on the asymmetric behavior of membranes in the cell (7, 8, 9).

The enzymes which control the cellular processes are themselves specified by the genetic apparatus of the cell from information stored in the cell's DNA (10). Modification of the cell's structure can be regulated by the relative rate of synthesis of each possible enzyme. A mechanism for this regulation has been worked out for bacterial cells (11). The extent to which it exists in the cells of multicellular organisms is still uncertain.

Periodicities are found in many cellular processes. Some periodicities have been attributed to chemical reaction systems which are so coupled as to oscillate (12, 13). While the life cycle periodicity of the bacterial cell has been attributed to sequential activation of the cistrons along the single DNA molecule making up its single chromosome, there is no good evidence for this. Even if that mechanism were involved in the life cycle of cells of multicellular organisms, it would be complicated by the presence of multiple chromosomes in each such cell.

Daily, tidal, lunar, and annual cycles of cell activity also have been observed (14). These cycles have a different mechanism than the life cycle, since they are relatively independent of temperature; an exception is the mitotic cycle , which has Q_{10} values of 2 to 3 (14, 15). The mechanisms of these cycles are not known. An extensive theoretical consideration has

been given the periodicities possible in a resting cell by both strong and weak couplings among all the biochemical processes in the cell (16). However, the model developed is limited in that it does not permit quantitative predictions, does not apply to growing cells, and (most importantly) does not switch the system from one semi-stable state (mode) to others as the environment changes.

2. Intercellular Communication

The most obvious way for cells to affect each other is by the passage of molecules or smaller particles from one cell to another. This could allow a cell which lacks some of the enzymes needed to supply the energy for completing a metabolic synthesis to obtain the energy via products from a neighboring cell which contains the missing enzymes, and thus to carry out the complete metabolic synthesis. This has been shown to occur for mesodermal (stromal) cells of the cornea, which, because of lack of some oxidative enzymes oxidize glucose only to lactic acid. The lactic acid diffuses to the neighboring epithelial cells and is there oxidized further. This generates a flow of electrons from stroma to epithelium, which flow is a source of energy for transport of fluid or ions between those tissues. At least two processes of molecular synthesis in the stroma have been shown to depend on close juxtaposition of the epithelium (17).

In the developing embryonic cells of animals, it has been shown that the differentiation of pancreatic epithelium cells to become secretory cells requires that they be growing in close proximity to mesenchyme cells (18). This is also needed for the differentiation of epithelial cells in the salivary gland, thyroid, skin, kidney and thymus (19). The differentiation of pancreatic epithelium has been shown to depend on the passage from the mesenchyme to the epithelium of materials which seem to be large molecules, are deactivated by trypsin, and have low mobility under physiological conditions (i.e., through 0.1 micron filter pores) (18).

It is known that hormones act, at least in some cases, by altering the metabolic activities of the target cells. This has been extensively studied in a line of cultured hepatoma cells. These cells show an increased content of the liver enzyme tyrosine aminotransferase within 15 minutes after exposure to adrenal steroids. After 6 to 12 hours' exposure, the content rises to a new steady state which is 5 to 15 times higher than the basal activity. A model has been proposed to explain these and other data (19). It assumes that the cell's DNA contains a structural gene, which specifies the induced enzyme, and a regulatory gene. During a large portion (the inducible phase) of the cell's life cycle the structural gene is transcribed continuously into messenger RNA, which is then translated into the enzyme. During the inducible phase the regulatory gene, also, produces its protein product, the repressor, at a constant rate. The repressor and its messenger RNA are assumed to be very labile. The repressor is assumed to reversibly inhibit the translation of the structural gene's messenger RNA into the enzyme. That messenger RNA can be degraded only when so inhibited. Finally, it is assumed that the action of the repressor is directly or indirectly prevented by the presence of

the hormonal inducer. Therefore, in the presence of an inducer, the repressor is inactivated, messenger translation and synthesis of the enzyme occurs, and the degradation of the messenger RNA is prevented. Thus, the concentration of the messenger RNA increases, because the transcription of it from the DNA is continuous.

During the noninducible portions of the cell's life cycle, the transcription of both the structural and regulatory genes is assumed to be repressed by a process which is insensitive to the inducer. Under these conditions, translation of the pre-existing messenger RNA continues, and because of the absence of the repressor, the messenger RNA is more stable than during the inducible phase. Obviously, in these portions of the life cycle the hormone has no effect on the system.

The mechanisms controlling mutual adhesion and aggregation of embryonic cells represent another aspect of intercellular communication. These have been studied by disassociating the cells of embryonic tissues, suspending selections of them in culture media, and allowing them to reaggregate. It was shown that the progressive reattachment and aggregation of these cells depends on unimpaired protein synthesis, and is reversibly inhibited by ribonuclease. The proteins and RNA involved appear to show a rapid turnover (20). The adhesivity and other surface properties of embryonic cells appear to be affected by neighboring tissues so that the differentiating cells move to particular alignments and shapes of distribution, thus forming new tissues (21). The effects of the neighboring tissues show polarities and gradients that depend on their position in the embryo (21, 22).

There are some features of differentiation in higher cell types which call for a more elaborate mechanism than that of bacteria. The most striking of these features are: (a) the change in state of differentiation is often mediated by a simple chemical signal, such as a hormone. (b) a given state of differentiation tends to require the integrated activation of a very large number of noncontiguous genes. This has led to the proposal of a new set of regulatory mechanisms for higher cells such that multiple changes in gene activity can result from a single initiatory event (27). The specific suggestion is that the inducer molecule activates a sensor gene which in turn activates an integration gene whose function is the synthesis of activator RNA. This RNA acts on several regulatory genes each linked to a structural gene to permit synthesis of the several enzymes specified by those structural genes.

More recent studies (23, 24) make clear that the formation of aggregates of cells involves not one process, but many. Therefore, the effects of various chemicals, such as protein synthesis inhibitors, depend on the particular conditions of the cell suspension. For example, horse serum in the suspension medium contains a macroglobulin-like protein that promotes the aggregation of neural cells of the retina and a second factor that promotes the aggregation of limb bud cells; while chick and calf sera contain factors that inhibit the aggregation of neural retina cells (24). It is possible that the last step in the formation of tissues involves the secretion of a fibrillar intercellular material which binds the cells together (25).

(a) Intercellular contacts and communication. - It has been found that in certain tissues there is a low electrical resistance between neighboring cells, although high resistance to the exterior of the cells (28, 29). In salivary gland and renal epithelia virtually all the cells of the entire epithelium were so connected into a single system (29). In bladder and sensory epithelial each cell was coupled to a chain or a small group of cells, with perhaps several such coupled systems in parallel (29). Molecules of up to 69,000 molecular weight (36 Å equivalent hydrodynamic radius) diffused freely between such electrically coupled cells, though molecules of weight 127,000 did not (30).

The junctions by which these cells are coupled can be disconnected by removing calcium ions with a chelating agent in the external medium; each cell then sealing itself off irreversibly as a unit (31). In the salivary gland of Drosophila the only structure which seems a likely site for the coupling is the septate junction, at which the opposing cell membranes are joined by a honeycombed bridge (32). The electron microscopic appearance of the septate junction is unchanged when the cells are uncoupled by calcium chelation (33). The coupling structure is thought to consist of the honeycomb of tubes connecting the opposed cells. The tube walls act as the perijunctional insulation from the extracellular space. Where each tube abuts onto the membranes of the two cells, those junctional membranes are highly permeable. The uncoupling of such a structure has been shown to occur in two steps (34). First, the reduced extracellular concentration of calcium ions renders the interior of the coupled cell system permeable to calcium ions, presumably chiefly at the perijunctional insulation. Then, the extracellular calcium ions enter the interior (including the cell cytoplasm) and cause the permeability of the junctional membranes to fall. This can be demonstrated even if the external calcium ion concentration is left high enough (10^{-4} M) not to affect the adhesion of the cells. It is hypothesized that the junctional membrane becomes highly permeable only when, as is normally the case, the concentration of free calcium ions in the cytoplasm is very low (35).

It has been shown with separated sponge cells that these high permeability junctions form within minutes in any regions of the cells' surface membranes which are brought into contact (36). The junctional membrane permeability decreases at least three orders of magnitude when calcium or magnesium ion activities in the cell interior are varied from the normal value of 10^{-5} M to 10^{-3} M, such as normally prevails in the extracellular medium (37). Maintenance of the low interior activities depends on (a) maintenance of the perijunctional insulation (35), and (b) maintenance of cellular metabolism, presumably metabolism supporting calcium and magnesium ion extrusion (38). Ouabain, which is a specific inhibitor of sodium - and potassium - activated ATPase and of sodium extrusion, produced depolarization but not junctional uncoupling (38).

Embryonic cells (blastomeres) in the morula stage are coupled by high permeability junctions (39). These junctions form between the daughter cells during cleavage of a blastomere. If two isolated blastomeres are brought into contact, such a coupling is formed in what appears to be only seconds

(39). This includes formation of the perijunctional insulation and change of the junctional membranes from low to high permeability.

(b) Communication through cell junctions: its possible roles in growth control and differentiation. - The results and implications of the studies of the intercellular junctional couplings are well covered in a review paper (40). These junctions have been found in salivary glands, neuroglia, renal tubules, certain sensory epithelia, urinary bladder, liver, skin, sponges, brown fat, thyroid, stomach epithelium, and embryonic tissues. It has been shown in salivary gland cells of Drosophila larvae that particles of about 3×10^3 molecular weight can pass through the junction; and dyes of lesser weight have been shown to pass through junctions in squid embryos, in certain crayfish axons, and in certain tissue-cultured fibroblasts.

These coupling junctions seem to form at the points of contact of adhering cells (of certain types) if the intracellular calcium concentrations are at their normal low levels (below 10^{-5} M) and the extracellular calcium concentration at the normal value of about 10^{-3} M.

Many metabolites, hormones, and other cellular substances fall within the size range of particles which can pass through the junctional membranes. This probably includes substances which regulate gene activity. There is a difference between such control via the junctional couplings and via the hormonal systems. The latter emit control substances into the open extracellular spaces; the flow will have extremely poor spatial resolution. However, flow via the junctional system will be limited to only the group, chain or network of cells which are junctionally coupled. This would enable it to evoke patterns of differentiation with a spatial precision of a few cell diameters or less. It is noteworthy that junctional connection is extensively present in embryonic tissues. It is also noteworthy that cancerous cells of a type that normally has junctional communication show such junctions to be altered or absent (40, 41, 42). Indeed, the loss of these junctions may be an etiological factor of cancer.

Functional coupling may also provide an explanation of earlier experimental findings on the differentiation of amphibian ectoderm into neural tissue (neuralization). A variety of unspecific and seemingly unrelated agents were found able to initiate neuralization: removal of external calcium and magnesium ions, mechanically or chemically induced cell injury, anisotonicity, alkali, methylene blue, and various metabolic inhibitors. These agents have one thing in common: they are uncouplers of junctional communication. The first four agents act by opening the surface barrier to the influx of extracellular calcium and magnesium ions; and the last two, by interfering with calcium ion extrusion through nonjunctional membranes. Lithium ions are known to interfere with embryonic differentiation; they, also, have been shown to produce junctional uncoupling (43).

Models have been suggested (40) by which the junctional communication could regulate growth. The size of cytoplasmic volume joined by junctions could be signalled by the concentration of a diffusible substance produced by a cell. The position of a cell in a junctional system could be signalled

by the rate of decrease of the concentration of a substance it produces in a pulse, this substance diffusing faster through the junctional than the non-junctional membranes, and thus decreasing more rapidly for centrally located cells.

(c) Detailed models for the achievement of spatial patterns in cellular differentiation. - An analysis has been made of the general requirements for any mechanism which can specify spatial patterns of cellular differentiation (44). Such a mechanism must show pattern regulation, which is the ability of the system to form the pattern even when parts are removed or added, and to show size invariance. The general requirements are: "(1) There are mechanisms whereby cells in a developing system may have their position specified with respect to one or more points in the system. When cells have their positional information specified with respect to the same set of points, this constitutes a field. (2) Positional information largely determines, with respect to the cell genome and developmental history, the nature of the molecular differentiation that the cell will undergo. The general process whereby positional information leads to a particular cellular activity or molecular differentiation will be termed the interpretation of the positional information. The specification of positional information in general precedes and is independent of molecular differentiation. (3) Polarity may be defined in relation to the points with respect to which a cell's position is being specified: it is the direction in which positional information is specified or measured. (4) Positional information may be universal; that is, the same mechanisms that specify positional information may be operative in different fields within the same organism as well as in quite different organisms from different genera or even phyla. (5) The classical classes of pattern regulation, whether in development or in regeneration ... are largely dependent on the ability of the cells to change their positional information in an appropriate manner and to be able to interpret this change." These general concepts were shown to fit the details of development of the sea urchin embryo, regeneration in hydroids, pattern formation in the insect epidermis, and the development of the chick limb.

More specifically, the mechanism is required to specify for about ten hours the position of about 50 cells in a line. This might be done by the low ionic resistance junctions discussed in the preceding sections. However, there is experimental evidence that, at least for polarity potential, cell to cell contact is not necessary (45). A more specific mechanism for producing spatial and temporal organization in differentiating tissues has been proposed (46). It assumes that individual cells produce periodic signals as a result of internal periodic biochemical events. (Possible mechanisms of biochemical oscillations are considered in references (12, 13) and (16)). These signals act intercellularly so that the fastest cells entrain all cells in the tissue. (Entrainment of loosely coupled oscillators is considered in reference (47)). A periodic event is then propagated outward from the pacemaker region, synchronizing the tissue. A refractory period is necessary to guarantee unidirectional propagation. The resulting wave-like propagation of activity can be described by a wave equation and a set of boundary conditions. An underlying gradient of frequency of the event establishes the relative position of the pacemaker region and the direction of propagation. Posi-

tional information is provided by a second event which propagates more slowly than the first, the information being the phase difference between the events. A third event is used to regulate the pattern of phase difference and thus establish size invariance. The developmental axis (and polarity) is defined as the longest trajectory orthogonal to the surfaces of constant phase difference, beginning at the pacemaker region and ending at the regulating region. In a similar way, information can be specified along other axes.

This model meets the requirements of the general model described above, and is shown to fit data on development and regeneration in hydra, and positional determinations in the early amphibian embryo. The model is also developed in some detail to account for the specific neural mapping of the retina onto the tectum in amphibians. (This mapping will be discussed in more detail in a subsequent section.) It is worth noting that junctional coupling has been found between cells in the neural plate, neural fold, and epidermis of neurulae in the amphibian embryo (39).

3. Information Input to the Nervous System

The first requirement is that the organism must receive information about its environment. In physical terms, this information must be collected by the interaction of atoms or molecules or quanta of radiation from the environment with molecules or atoms of the organism. The amount of such information reaching the organism's surface is enormous. Consider yourself sitting in a well-lighted room. Every square centimeter of your exposed skin is receiving about 10^{14} photons every second from the environment. That exposed skin is also being hit by about 10^{25} molecules per second on every square centimeter. If every one of these collisions were to be reported to the organism, the result would be chaos - a flood of information much too great to be handled. Therefore, an important function of the information collecting devices of an organism is to select those quantal events or patterns of events which are highly relevant to the survival of the organism.

(a) Sensory receptors. - Those cells of the organism whose function is to collect information about the interface between organism and environment and pass this information on to other cells are known as "sensory receptor cells". These cells often act as transducers, that is, they change the form of energy with which the environment acts on them to another form of energy which can be transmitted to other cells. These cells also amplify the energy involved, using energy sources from within the cell and the organism. This transduction takes place either at a specialized membrane of a nerve cell (48) or at a specialized ciliated cell (49).

Let us first consider the nerve membrane type of transducer. A typical and well-studied example is the Pacinian corpuscle, which is a touch receptor found just under the skin. When the bared nerve ending is stimulated by pressing on its surface with a tiny rod, the membrane potential decreases (50). During a maintained mechanical stimulation, there seems to be a constant net influx of positive charge to the nerve ending. It is known that various ions such as sodium are present in a lower concentration inside the nerve ending

and rest of the cell than in the surrounding liquid medium; whereas other ions such as chloride and potassium are present in a higher concentration inside the nerve ending than in the surrounding liquid medium. Normally, the cell membrane acts as an insulator which prevents the flow of these ions into or out of the cell. The observed change in membrane potential caused by mechanical stimulation (called the generator potential) is thought to result from the following mechanism. The mechanical distortion of the membrane causes the membrane to become more permeable to some ions (probably all the small ions) and the resulting ion flow along the ion concentration gradient tends to move the membrane potential toward the combined equilibrium potential of zero. When the mechanical distortion of the nerve ending ceases, the membrane potential returns exponentially toward the resting level, as would be expected if the cell membrane behaved as a leaky insulator (parallel capacitance and resistance) and thus the whole dendrite could be considered as a leaky cable. It has been shown that each mechanically distorted region of the nerve ending membrane acts independently (50, 51).

Most sensory receptors based on specialized nerve membranes seem to act in the same way; that is, stimulation acts on the nerve membrane so that its permeability to ions is changed. It has been found that direct mechanical distortion of an ordinary nerve cell fiber will produce such a change. Such changes can also be produced by chemical reactions directly with the nerve membrane, as happens at the chemical synapse between nerve cells. Permeability changes can be produced by a chemical change which is initiated by the absorption of light, as happens in photoreceptor cells. And, there are some nerve cell membranes (such as the axonal membrane) whose structure can be distorted by a change of the electric potential gradient across it, with a resulting change in its permeability to ions. It has been found that each stimulated unit area of membrane acts both as a source of electric charge flow (ion flow) and as a shunt to the currents from all the other stimulated unit areas. Therefore, as there is an increase in the number of stimulated units of membrane (or stimulated area), the resulting potential change does not increase linearly with the area, but at a diminishing rate. The peak voltage change, V, can be described by the following equation:

$$V = k \frac{Ebx}{1 + bx} \qquad \text{(Eq. 1)}$$

where x is the fraction of units excited or the ratio of the stimulated excitable area to the total excitable area, E is the resting membrane potential, $1 + b$ is the ratio of E to the membrane potential reached with maximum stimulation, and k is a constant which depends on the location of the electrode.

Note that the size of the x is a measure of the intensity of the stimulus, and that the resulting potential is proportional to the stimulus intensity divided by the quantity (one plus some constant) times the stimulus intensity. If such an equation is plotted as voltage versus w, the logarithm of the stimulus intensity, it can be shown (52) that

$$V = 1/2 + 1/2 \tanh (w - w_o). \qquad \text{(Eq. 2)}$$

Note that there is a large range of voltages which lie along a straight line function of the logarithm of stimulus intensity. Biologists have often described the response of the cell or even of an organism as being proportional to the logarithm of the stimulus intensity, and speak of a logarithmic transform of stimulus intensity. Equation 2 above is a more exact description of the relationship, since it covers the entire range of stimulus intensity, whereas the logarithmic relation applies to only one part of the range.

It has been shown that in at least one receptor, the light receptor of the Limulus lateral eye, the change in membrane potential upon illumination results from an illumination-evoked decrease in the active ionic transport occurring at some portion of the cellular membrane (53, 54).

The second general type of sensory receptor is the ciliated cell. One common type is the "hair cell", which responds to bending of its cilia. The bending causes a change in the permeability of the cell membrane to certain ions, with a resulting change in potential across the cell's membrane. It is known that the hair cells are asymmetric in structure, and there is reason to believe that bending of the cilia in one direction along the axis of asymmetry results in an increase of the cell's membrane potential while bending in the reverse direction leads to a decrease in the potential (55). Such hair cells are used to detect movement of the organism with respect to its environment in the lateral line organs of the fish and in the semi-circular canals of mammals. In the electric fish, some of these hair cells have become specialized to detect the small electric potential gradients resulting from the discharge of electric pulses by the fish into the surrounding water. This is used for the detection of obstacles by the distortion the obstacles introduce into the electric potential gradient around the fish, and also used for communication between electric fish (56). Similar hair cells are used with the very elaborate accessory structure of the ear to detect oscillations in air pressure (sound) (57).

4. Transfer of Information Along a Nerve Cell

(a) The general form of the nerve cell. - Because nerve cells are used for the distribution of information to other cells throughout the organism, and because the distribution may involve branching and multiple connections, there has been some confusion between the anatomical and functional relations of the various parts of the nerve cells. These relations are clarified in two papers (58, 59) to which the reader is referred. The suggestion is made in the first of these papers that those regions of the nerve cell in which a generator potential is produced and through which it is passively conducted be referred to as the dendritic zone of the cell; that those regions which propagate an action potential (see below) be referred to as the axonal zone of the cell; and that those regions at which the cell makes contact with other cells and transfers information to them be referred to as the telodendritic zone. In general the dendritic zone is the input portion of the nerve cell and the telodendritic zone is the output portion.

(b) Electronic conduction. - As mentioned previously, the membrane of every cell acts like a leaky capacitor. Therefore, a current flow through the membrane at any point on the cell will result in a change in potential across the membrane throughout the cell; this potential change drops off with distance in a manner that depends upon the geometry of the cell. This is known as passive, or electrotonic, conduction. It has been claimed that electrotonic conduction is used by some nerve cells to transmit information over distances as great as two millimeters, however, this mode of conduction is normally not used for distances greater than about 200 microns. For greater distances, the amplitude of the electrotonic potential is at some nearby point in the cell (60) converted to frequency of firing of nerve impulses, which in turn are propagated as all-or-none (digital) pulses over the greater distances. The propagation velocity varies from 0.02 to 1.0 m/sec. and is larger in the larger diameter axons.

The change in membrane potential at the transducer nerve ending is called the generator potential because the associated current flow spreads to the axon and leads to the generation of nerve impulses. The generator potential, in contradistinction to the action potential which accompanies the nerve impulse, can vary in amplitude and is therefore called a graded potential. Its amplitude usually depends upon the stimulus intensity in accord with the relation given above (Eq. 1).

(c) Axonal, or all-or-none conduction. - The net positive charge that flows in through the stimulated membrane of the receptor nerve cell spreads along that membrane as though along a leaky (capacitative) cable, decreasing exponentially with distance from its source. Some of this current flows out through the nearest gap (node of Ranvier) in the myelin sheath of the nerve axon. Such gaps or nodes are spaced at intervals of about 2 mm along the myelinated nerve fiber. The membrane at that node is of the type called electrically excitable in that its permeability is altered by a decrease in the electric gradient across it so that it becomes more permeable to sodium ions. The sodium ions flow in from the outside, bringing a positive charge with them and dropping the electric gradient across the membrane still more, and thus by positive feedback causing the membrane to reach its maximum permeability to sodium ions and causing the potential across the membrane to go positive inside by a maximum amount. During this process, there is a delayed increase in the potassium permeability of the membrane which then allows an increased potassium ion flow that acts to restore the voltage gradient across the membrane and shuts off the sodium ion flow within one millisecond. Several milliseconds are required for the potassium permeability to return to normal.

This entire course of events is called the nerve impulse, and the accompanying change in membrane potential is called the action potential or spike potential and is an all-or-none response in the sense that it goes to the maximum of which that membrane is capable at the time. During the existence of an action potential at a node, the positive charge entering at that node flows out through neighboring nodes, decreasing the membrane potential there and in turn triggering those nodes to release the stored energy of the ion concentration gradients to produce an action potential. In this way, the

action potential is transferred from node to node along the axon and is thus conducted through the organism to reach other cells. This form of transfer is termed active (energy-added) conduction. This mechanism for information transfer involves a considerable safety factor against the dropout of impulses - the amplitude of the action potential at any one node along the axon can be decreased to as low as 1/5 of its normal value without affecting the transfer of the pulse to the next node, where it will be produced in its normal full value.

After firing an action potential, the axonal membrane must recover its normal low permeability to potassium before it is capable of firing a second action potential. This period of lessened tendency to fire, upon decrease of the potential across the membrane is called the refractory period. The time interval between action potentials is decreased if the applied generator potential, and thus the decrease in membrane potential, is increased. It has been shown that the rate of firing of action potentials is a direct linear function of the amplitude of the generator potential (61). This relation has also been tested by using an external current source to put positive charge into a cell via an electrode contacting the inside of the cell and observing that the change in frequency of the firing was a linear function of the applied current (62). Since the usual generator potential varies with stimulus intensity according to Equation 1, and since that equation can be approximated for a wide range of stimulus intensities by a logarithmic function, it is not surprising that the frequency of firing of nerve cells attached to a receptor has often been observed to be approximately a logarithmic function of the stimulus intensity applied to the receptor.

The rate of propagation of the action potential along an axon (conduction velocity) is directly proportional to the leakage resistance of the axonal membrane and is inversely proportional to the internal resistance along the length of the axon. Since the latter is inversely proportional to the cross-sectional area of the axon, the conduction velocity is proportional to the square root of the axon's diameter. Typical conduction velocities are from 2 m/sec to 100 m/sec in myelinated axons and 2 m/sec or less in the usual unmyelinated axons, although it is 20 m/sec in the giant (500 micron diameter) axon of the squid.

5. Encoders

The anatomically specialized regions of contact between cells which serve the function of rapid transmission of information from cell to cell are known as the synapses. There are two general types of synapses; electrical and chemical. These synapses introduce their own transfer function into the relation between the output of the first cell (presynaptic cell) and the input of the second cell (postsynaptic cell). These transfer functions come in a great variety, and are responsible for much of the information processing (encoding) that occurs in the nervous system (63).

(a) Electrical synapses. - One type is the unpolarized electrical synapse in which the membranes of the two cells are fused together in the

region of the synapse, which then acts as a low impedance connection between the interiors of the two cells (an ephaptic junction). This simply allows an electrotonic spread of potential from the presynaptic to the postsynaptic cell, multiplied by a constant which depends upon the relative geometry of the synaptic area. This particular type of electrically excitable synapse is unpolarized, in the sense that current flows and voltage changes are passed equally well from one cell to the other. If the non-synaptic membranes of the two cells are electrically excitable, then transmission of action potentials occurs in either direction along the two cells, just as it does in an axon. The junctional couplings described in Section 2 are of this type.

A second type is the polarized electrical synapse in which the membranes at the ephaptic junction show different resistances to different directions of current flow. Thus, the conductance in the direction of flow tending to reduce the membrane potential of the post-junctional cell may be twenty times greater than that in the opposite direction (63). These polarized ephaptic junctions allow conduction of action potentials in only one direction.

(b) Chemical synapses. - The second, and by far the most common type of synapse is the "chemical transmitter synapse". Anatomically the synapse consists of a region in which specialized portions of the membranes of the two cells lie parallel to each other with a gap of about 300 Å between them. Reduction of the membrane potential in the telodendritic terminal (presynaptic terminal) somehow causes the release from that terminal of packets of chemical (the transmitter substance) into the space between the two cells. These chemicals diffuse across a synaptic space (synaptic cleft) and interact with the postsynaptic membrane to increase its permeability to certain ions.

(1) Excitatory synapse. - A synapse is called excitatory if its activation tends to cause the postsynaptic cell to have a decreased negativity of its membrane potential (a depolarization) and an increased tendency to fire action potentials. In this type the chemical transmitter usually causes an increased permeability to sodium ions at the postsynaptic membrane. The inward flow of sodium ions along their electrochemical gradient leads to depolarization. This, in turn, increases the excitability of the postsynaptic cell's electrically excitable membrane and increases the tendency to fire action potentials (63).

(2) Inhibitory synapse. - A synapse which upon activation tends to reduce the firing of action potentials by the postsynaptic cell is called inhibitory. Here the chemical transmitter usually causes an increased permeability to chloride ions or potassium ions or both at the postsynaptic membrane. The resulting ion flow may be either inward or outward depending on the difference between the membrane potential and the electrochemical gradient (which usually equals the equilibrium potential for its concentration gradient) of the ion. The result may be either a hyperpolarization (increase in membrane potential) or a depolarization of the postsynaptic cell. The ion flow tends to clamp the cell's membrane potential at the equilibrium potential of the ions involved, and tends to prevent depolarization of the membrane in response to activation of excitatory synapses. It also tends to prevent the positive feedback depolarization that occurs at the axonal membrane during

production of an action potential. The net result is a reduced firing of action potentials by the postsynaptic cell (63).

Frequently, at a synapse there is a spontaneous, random release of the chemical transmitter packets (63). The postsynaptic potential change (or generator potential) evoked by a single packet is quite small, being only 80 microvolts in the nerve-muscle junction of the crayfish (64). These miniature generator potentials might be considered to be a source of noise except that in nerve axons and in muscle cells there is a certain amplitude of potential change which must be exceeded before an action potential is initiated. This threshold potential change may be as much as 200 times the potential change caused by a single chemical packet. Thus, there is a very large safety feature (noise discriminator) built in to discriminate against these spontaneous miniature postsynaptic potentials (64).

(c) Determination of synaptic action by the postsynaptic membrane. - The action at each chemical synapse depends more upon the nature of the postsynaptic membrane than upon the particular chemical involved. It seems to be a general finding that any given neuron will release the same chemical transmitter at all of its telodendritic terminals. However, it has been demonstrated that this transmitter may have an excitatory effect at one terminal (cause a depolarization of the postsynaptic cell) and have an inhibitory effect at its terminal on another cell (cause hyperpolarization of the postsynaptic cell) (65).

It is also possible that at a single synapse the telodendritic terminal may release two different chemicals which both act as excitatory inputs at the same synapse, but with different sensitivities and time constants of action (66). This implies that the postsynaptic membrane at this synapse contains patches of different kinds of membrane, each sensitive to one of the transmitter chemicals.

As a further possible complication, it has been found for some cells that a single synapse may be either excitatory or inhibitory depending upon the rate of release of its chemical packets (67). In this case, the chemical packets are all thought to be the same (acetylcholine). The postsynaptic terminal is thought to have two types of membrane, both sensitive to acetylcholine. One type has a low threshold for activation and produces an excitatory postsynaptic potential. This type is readily blocked by higher concentrations of acetylcholine, and then the other type of receptor membrane, which has a high threshold, comes into play and evokes inhibitory postsynaptic potentials.

The multipurpose synapses discussed in the two preceding paragraphs have been found so far only in the mollusc, Aplysia. Since these molluscs have only about 2,000 or so neurons, they need to make each neuron serve as many functions as possible. This may be the reason for the multipurpose synapses, and it is less likely that each synapses will be found in vertebrates, which have anywhere from one to ten billion nerve cells.

(d) One-to-one transmission. - A very simple chemical synaptic arrangement between two large nerve terminals, the giant synapse of the squid, has been studied to determine the input-output relation of the synapse. This was done by placing an electrode in the postsynaptic terminal to determine its depolarization and simultaneously placing two electrodes in the presynaptic terminal, one to inject a current to change its membrane potential and the other to measure that membrane potential (68). The cells were immersed in a solution of tetrodotoxin, which has the property of preventing the generation of action potentials by an electrically excitable membrane. This permitted the experimenters to measure slow changes in membrane potential at the presynaptic and postsynaptic terminals without interference by the rapid changes of the action potential. Their findings were that the curve relating the input and output, that is, the presynaptic depolarization to the postsynaptic depolarization, fitted very well to the curve of Equation 1, except for the low values where a Poisson curve fitted better. In this particular synapse the action potential normally can invade the presynaptic terminal, and the depolarization accompanying an action potential equals the presynaptic depolarization. Therefore, each action potential entering the presynaptic terminal produced a very large (near maximum) depolarization of the postsynaptic terminal, and this, in turn, is enough to generate an action potential in the postsynaptic cell. So this particular synapse acts simply as a one-to-one transmitter of action potentials from the first cell to the second.

(e) Fatigue. - If a synapse is stimulated at a rapid repetition rate for several seconds, the postsynaptic potentials are observed to decline in size, and this has been traced to a decreased number of chemical packets released for a given depolarization of the presynaptic terminal. The probable mechanism of this phenomenon, which is called "fatigue", is that the number of chemical packets in the presynaptic terminal has been reduced faster than new ones can be synthesized (64).

(f) Facilitation. - For a number of varieties of chemical synapses, it has been found that moderately intense, repetitive depolarization of the telodendritic terminal will cause an increase with time in the rate at which chemical packets are released from the terminal into the synapse for a given depolarization at the terminal. This leads to the production of increasing amplitudes of postsynaptic depolarizations during a series of fixed-amplitude postsynaptic depolarizations. Such a phenomenon is called "facilitation" (64).

It has been noticed that after transmission of a short burst of high frequency action potentials to the presynaptic terminal, there is a facilitation of its effectiveness which may last for 60 seconds or more. Immediately after the burst, the frequency of the spontaneous miniature postsynaptic potentials increases for several seconds and then decreases for a period of up to four minutes (64). (It is possible to surmise that these phenomena can be explained by the following model. First, assume that there are two mechanisms by which chemical packets can be released from the presynaptic terminal: one causes a spontaneous release of the packets, the other causes release when the terminal is depolarized. Second, assume that the spontaneous release

mechanism takes a significant period of time to operate, so that a number of chemical packets are involved in that mechanism at any instant of time. Third, assume that during the stimulating burst and during the following few seconds the same process that causes facilitation of the depolarization release mechanism also causes the spontaneous release mechanism to rapidly free into the synaptic cleft all the packets involved in the latter mechanism. There must then be a wait of several minutes before fresh packets enter and pass through the spontaneous release mechanism.)

(g) Delay. - The minimum time that has been observed between depolarization of the presynaptic terminal of a chemical synapse (at a motoneuron) and the appearance of the postsynaptic potential is 0.5 msec. With weaker presynaptic depolarization, this delay can increase to 1.0 msec (69, pp. 303-306).

(h) Temporal integration. - At some types of excitatory synapse, the chemical transmitter continues to act on the postsynaptic membrane for periods of 15 msec or more, instead of the usual 1 to 2 msec. If additional action potentials arrive at the synapse during this period, they cause an increased change in the postsynaptic potential and an increased activity of the postsynaptic cell. The transfer function of such a synapse would include an integrative time constant of about 15 msec.

The cells of the lateral geniculate body, which are the cells that relay signals from the eye to the cerebral cortex, have been shown by transfer function analysis to have an integrative time constant for their input synapses of about 16 msec (70). This suggests the integrative mechanism mentioned in the first paragraph.

(i) Effects of synaptic location and cell shape. - The consequences of different geometrical locations and arrangements of synapses on the cell have been studied (71). Mammalian motoneurons have a high density of synaptic connections over the surfaces of their dendrites and their cell bodies. The motoneuron's cell membranes are electrically excitable at the cell body and the base of the dendrites and electrically inexcitable over the remainder of the dendrites. The dendrites can be treated as cylindrical leaky cables, and conduct the postsynaptic potential change passively. The potential change decreases with distance as it is conducted, the rate of decrease being greater in the smaller diameter branches of the dendrites. For the typical branched dendrite, the postsynaptic potential originating at a peripheral terminal (i.e., one farthest from the cell body) is reduced to $^1/e$ to $^1/e^2$ of its value at the synapse during its conduction to the cell body. Those synapses acting directly on the spherical cell body cause a membrane voltage change that is everywhere on the cell body very close to that at the postsynaptic terminal. Although the individual brief postsynaptic potentials at the dendritic peripheral terminals are significantly attenuated in their spread to the soma, the very large number of dendritic synapses (of the order of 10^3 to 10^4) permits a ceaseless spatiotemporal spattering of synaptic events, the integration of which can provide a finely graded background level of depolarization (a biasing of neuronal excitatory state). The sluggish transient characteristic of dendritic electronic conduction provides for a smoothing at

the cell body of the temporal input pattern received at the dendritic periphery. If the excitatory impulses are delivered to synapses on separate dendritic peripheral terminals, there will be approximately linear summation of their effects at the cell body. However, when the excitatory impulses are delivered to synapses located close together on the same dendrite or on the cell body, the summation will not be linear, but more like that of Equation 1. Note that the cell body acts as the site of evaluation of the synaptic inputs and the operations performed on them. The output of the motoneuron is the firing rate of its axon.

(j) Spatial summation. - While stimulation of one, or even a few, telodendritic terminals synapsing on a given cell may not be sufficient to give an excitatory postsynaptic potential that exceeds the depolarization threshold for action potential initiation in that cell, the excitation of a larger number of telodendritic terminals would cause a sufficiently large postsynaptic depolarization to initiate action potential production. Thus, the postsynaptic cell is acting as a coincidence counter for activity in a number of telodendritic terminals. This behavior is termed spatial summation. This summation could be either of the two types described in the preceding paragraph.

(k) Divisive inhibition, postsynaptic. - The effectiveness of the excitatory inputs to a neuron can be reduced (inhibited) by two mechanisms. The first, postsynaptic inhibition, has been discussed in the section on inhibitory synapses. For particular geometries of the neuron the excitatory effectiveness is divided in proportion to this inhibition. Those geometries are ones which cause the excitatory potential source (e.g., excitatory synapses at dendritic terminals) to produce a current which has to pass through a relatively large resistance (e.g., that of a long dendrite) to reach the site of nerve impulse initiation (e.g., the axon hillock). The potential change, and corresponding frequency of firing, at that site is proportional to the current reaching the site. This current is reduced if a relatively low resistance shunt is placed across the neuronal membrane at the site side of the high resistance path. The reduction is approximately in proportion to the conductivity of the shunt. An inhibitory synapse can act as such a shunt if it acts to increase the membrane's conductivity to ions other than those involved in the initiation of the nerve impulse.

(l) Subtractive inhibition. - In certain specially structured cells, postsynaptic inhibition can be directly subtractive from excitation, instead of being divisive. A well-studied example is the eccentric cell of Limulus' lateral eye (72). That neuron receives excitation as depolarization produced at its single dendrite via passive electrical conduction from the light-excited photoreceptor cells (retinula cells) of that eye. The depolarization is passively conducted with little loss along the electrically inexcitable membranes of this large diameter dendrite and large cell body to an axon about 4 microns in diameter. For the first 200 microns past the cell body, the membranes of the axon are electrically inexcitable. Also connecting to the axon in this region are small branch fibers. Those closest to the cell body are outputs to synapses, those further away (more central) include inputs from synapses with branch axons from neighboring eccentric cells. The

conductance from inside to outside the cell provided by the membranes of the output branch fibers is 4 to 6 times that of the dendrite and cell body. It provides a constant shunt to the depolarization produced at the dendrite.

The synapses to the more central branch fibers are inhibitory and produce an increased conductance across the cell membrane. This increased conductance appears at the main axon and decays with a time constant of about 500 msec. Thus, input of action potentials to the inhibitory synapse at rates greater than 1/sec can be expected to produce summation of conductance increases to a nearly steady maintained level. This has been observed. The inhibitory input from neighboring eccentric cells is termed lateral inhibition. The increased conductance produced by the inhibition in this region of the axon shunts the depolarization reaching this region from the dendrite and reduces the depolarization passed still more centrally to the action-potential-producing site of the main axon. The frequency of firing produced there is linearly proportional to the depolarization at the site (about 3.2 impulses/sec/mv). The conductance increase at the main axon is closely constant for a given input of lateral inhibition despite any changes in the excitatory input and conductance at the dendrite. The effect is to produce a constant decrement in the firing frequency of the main axon. The final relation can be expressed (73) as:

$$r_A = e_A - K_{AB} (r_B - r^o_{AB}) \qquad \text{(Eq. 3)}$$

where subscripts A and B refer respectively to the eccentric cell being inhibited and the cell doing the inhibiting; r is the firing rate; e is the excitation measured as firing rate with no lateral inhibition present; K is a constant; and r^o_{AB} is the firing rate at which the neighboring cell, B, ceases to exert an inhibitory effect on the given cell, A. The equation applies only for $r_B \geq r^o_{AB}$. Note the direct subtraction of increments in the firing rate of cell B (times a constant) from the firing rate of cell A.

It is of interest that this cell also shows self-inhibition. There is an increased conductance at its distal axon each time it fires an action potential, with the same time constant of decay as for lateral inhibition (72). Apparently this is produced by transmission of the action potential via the closest of the branch fibers to inhibitory synapses with the more central branches. The self-inhibition produces those transients in the firing rate following changes of excitation which would be expected of a negative feedback with that time constant.

(m) Gating by presynaptic inhibition. - The second major mechanism of inhibition is presynaptic inhibition, which is any process by which a cell acts on the presynaptic terminal of an excitatory synapse of two other cells so as to reduce the effectiveness of that synapse (74). For example, the axon of a muscle stretch receptor in a muscle (a type Ia nerve fiber) might terminate in the spinal cord at an excitatory synapse on a motoneuron that controls contraction of that muscle. Axons of muscle stretch receptors in an antagonist muscle (e.g., flexor vs. extensor at the same joint) might terminate on and excite interneurons, which in turn excite a "D cell" in the spinal cord. The D-cell has an inhibitory synapse on the excitatory presyn-

aptic terminal of that type Ia fiber with the motoneuron. It is thought that normally an action potential conducted by that fiber will go all the way into the presynaptic terminal, depolarize it maximally, and cause a maximum release of chemical transmitter to the motoneuron. When the synapse of the D-cell on the presynaptic terminal is activated, the postsynaptic membrane of that synapse, which is part of the membrane of the presynaptic terminal of the type Ia fiber, becomes highly conductive to ions. This conductive area of membrane shorts out the action potential conducted along the type Ia fiber and prevents it from being conducted on into the presynaptic terminal (75). As a result, the terminal is not depolarized and does not activate the synapse with the motoneuron. Thus, when the D-cell is active, the synapse between type Ia fiber and motoneuron is "gated off".

(n) Divisive inhibition, presynaptic. - At some synapses the presynaptic terminal has electrically inexcitable membranes and so is not invaded by the conducted action potential. The depolarization of the terminal is only that which is passively conducted from the region of the axon where active conduction ceases (at the last part of the electrically excitable axonal membrane). At an inhibitory synapse on the presynaptic terminal, the ion conduction of the electrically inexcitable membrane on that terminal, postsynaptic to that synapse is increased by activation of that synapse. This region acts as a shunt to reduce the depolarization passively conducted to the presynaptic terminal and thus reduce the release of chemical transmitter from it. Such a mechanism has been described for presynaptic inhibition of the junction between the excitatory nerve fiber and the opener muscle of the crayfish's claw (76). The relation between activity sent to the inhibitory synapse and the consequent reduction of excitation sent via the inhibited synapse has not been measured. Presumably, it, like postsynaptic inhibition, would be approximately a division of the excitation by the inhibitory activity.

6. Categorization of Input Information - Pattern Recognition

The animal's task of adjusting to change in its environment is simplified if those changes are categorized according to some scheme of invariants in the environment. Thus, the object, "table", is an invariant even though the visual input from it will depend on the illumination and the angle and distance of viewing. Presumably, the nervous system arrives at these particular categories by trial and error during evolution, resulting in genetic determination of the schemes of categorization; and by trial and error during development of the individual, resulting in "learned" recognition.

(a) Pattern detection at receptor organs.

(1) Selectivity of receptor and accessory structures. - As has been discussed already, this selectivity allows only a very small part of the environmental changes to be signaled to the animal's nervous system. It already prepares a general categorization of those changes. Thus, visual stimuli arrive along one set of nerve fibers, touch, sound, smell, taste, etc., each along other sets.

(2) Categorization by spatial characteristics of a single cell is exemplified by those touch receptors in the skin that are the bare nerve endings of branches of a single nerve cell. The spatial spread of those branch endings differs from cell to cell. One cell may have them localized in a small region of the skin and thus signal only pressure on that small region. Another cell may have its branches widely scattered and could signal either strong stimulation of a single branch anywhere within the wide region so covered, or by the mechanism of recruitment (see below) could signal simultaneous weak pressure on many branches within that wide region.

(b) Categorization by simple interconnections.

(1) Recruitment is the name applied to the increase in number of active nerve fibers upon increase of stimulus strength. This depends on the inputs (receptor endings or synapses) of the fibers having a variety of thresholds (input strength necessary for activation) so that more are activated as the stimulus strength is increased.

(2) Convergence is the connection of many nerve cells or receptors, usually via synapses, onto a single cell. The activities of these many inputs can summate at the single cell. If these inputs are subject to recruitment, then increased stimulus strength can cause the input to the single cell to exceed its threshold and produce an output (e.g., cause a nerve cell to fire action potentials).

(3) Spatial localization by facilitatory convergence and divergence. - In a series of studies (77) it has been shown that there are about sixty touch receptors in the footpad of the cat and that these connect by axons from each receptor cell to an equal number of neurons in the dorsal horn of the spinal cord. Each of these spinal neurons (second-order neurons, i.e., second along the sensory input path) receives terminal synapses from several of the receptor cells (first-order neurons, i.e., first along the sensory input path). This is described as convergence of receptors onto the spinal neuron. They also receive synapses from other neurons. The pad was mechanically stimulated at a repetition rate of 10 per second under conditions which apparently cause no primary neuron to fire more than a single action potential at each stimulation. The number of primary neurons which fire was increased by increasing the mechanical displacement at the pad or by increasing the area of the pad being displaced by the mechanical stimulus. Each second-order neuron has associated with it a "receptive field", that is, an area on the pad in which (mechanical) stimulation could evoke a firing of the second-order neuron. Each of these receptive fields contained a central region in which the mechanical stimulus strength needed to just evoke a response was constant throughout. Peripheral to the central region the threshold rose as distance from the center increased. It was found that within the receptive field the excitation produced at the second-order neuron by a localized mechanical stimulus added to the excitation produced by a mechanical stimulus elsewhere in the receptive field, and these two excitations could sum over a period (summation period) of up to nine milliseconds. This summing of excitation could be explained as the stimulation of different first-order neurons all of which have excitatory synapses with the single

second-order neuron being observed. The interactions of two mechanical stimuli at different regions on the pad were found to be always excitatory to the spinal neuron and greater than the effect of either alone, such an action being termed facilitatory. In no significant number of the cases was the response found to be less than that to either of the stimuli alone; therefore, this rules out the possibility of the activity caused by one stimulus detracting from the activity caused by the other stimulus in a neighboring location (lateral inhibition). The question arises as to how a second-order neuron can get a localized, specific receptor field when only a single action potential was fired by each excited first-order neuron and when only excitatory activity was transmitted to the second-order neurons. The authors propose a model to explain this which depends upon: (a) The fact that the primary neurons had an assortment of thresholds to mechanical stimulation. (b) The fact that the primary neurons were spaced out fairly uniformly along the surface of the pad. (c) The fact that the mechanical stimulus produces a wave of mechanical deformation which travels laterally through the tissue of the pad at a known speed, and thus reached neighboring primary neurons with known delays, (d) the assumption that the output (action potentials) of each second-order neuron is sent to a number of neighboring second-order neurons. (e) The fact that a number of contiguous first-order neurons connect to a given second-order neuron, there being a topological ordering between the array of first-order neurons and the array of second-order neurons. (f) The assumption that six or seven action potentials must converge on a single second-order neuron within the summation period of a few milliseconds in order to cause that second-order neuron to fire. This model results in a pattern of second-order neuron firing which maps both the strength and the position on the pad of the mechanical stimulation.

(c) Collision of action potentials. - There is one type of first-order neuron, the external spiral neuron of the cochlea, which has many branches going to its single main fiber from hair cells spaced along the cochlea. It was predicted in 1960 (78) that if it was assumed (a) that each branch transmitted action potentials in the direction away from the hair cells, and (b) that at the branch junctions these action potentials were transmitted one way (toward the brain), then such a branched neuron would show an increase in firing rate with increasing stimulus intensity over a much larger range than would an unbranched neuron. A later theoretical paper (79) attempted to provide a physiological mechanism that would give the one-way transmission at the junction. The assumption was that the branch ran very close to the main fiber at the junction so that as an action potential moved from the more distant part of the main fiber toward the junction, current from it would pass through the branch and reduce the excitability of the nearby portion of the branch. Then, when the action potential reached the junction, it could not propagate backwards up the branch but only forwards on the main fiber. This theorist also pointed out that because of the increased diameter of the fiber at the junction point, the velocity of propagation of the action potential would be decreased and this would permit a closely following action potential coming along the fiber to catch up with it and merge into it so that only one action potential would be propagated further along the fiber. These assumptions lead to the same increase in dynamic range of the spiral nerve as had been predicted by the first paper. It was later found (80) that

some first order auditory nerve fibers do exhibit exactly that increased dynamic range as compared to the other types of auditory nerve fibers. While these findings are consistent with the proposed mechanisms, there is no direct evidence that such mechanisms do indeed exist in the cochlear nerve.

(d) Gating. - It has been observed that the output nerve cells of the retina, the ganglion cells, will fire a burst of action potentials when the eye is stimulated with a fixed duration flash of light. As the intensity of flash is decreased, the number of action potentials in the response also decreases. In some ganglion cells the number of action potentials will decrease down to one and finally zero; however, in some other ganglion cells, the smallest number of action potentials observed may be a burst of, say, four. As the flash intensity is decreased, this burst of four occurs for a smaller percentage of the flashes, the response to the other flashes being no firing at all.

This property of the minimum response of the nerve cell being a burst of action potentials could be explained on the assumption that the characteristics of the membranes of the cell body and axon are such that any depolarization of those membranes which is sufficient to initiate an action potential decays so slowly that several action potentials are fired during this time.

It is interesting to suggest the possibility of another mechanism, which is a form of gating. A recent electron microscopic study of the vertebrate retina showed (81) that the photoreceptors (rods and cones) contact and send excitation to the second-order nerve cells, the bipolar cells. The bipolar cells in turn synapse with dendrites of the third-order neurons, the ganglion cells, and send excitation to them. The synapses, however, are often of a peculiar organization (see Fig. 1) in which the synapse from the bipolar cell is to two side-by-side cell branches, one from the ganglion cell's dendrite and one from an amacrine cell. Just a short distance away, the same amacrine cell branch has a synapse that by its structure evidently acts from the amacrine cell onto that bipolar cell's synaptic terminal. It is reasonable to assume that the synapse from the bipolar cell to the ganglion cell's dendrite is excitatory. If it is further assumed that the synapse from the bipolar cell to the amacrine branch is inhibitory and that the synapse from the amacrine branch to the bipolar cell is also inhibitory, then the following behavior can be predicted. For a given level of activity at the amacrine branch, there will be produced a given level of hyperpolarization at the bipolar cell's synaptic region. As excitation (depolarization) is sent from the photoreceptor cells to the bipolar cell's synaptic region, this will be at first opposed by the hyperpolarization and will not result in any excitation being sent to the ganglion cell. At some critical level of incoming depolarization the bipolar's output synapse will begin to work and will send inhibition to the amacrine cell branch, which will decrease the hyperpolarization being sent from the amacrine cell branch to the bipolar cell's synaptic region. This will allow further inhibition by the bipolar cell of the amacrine cell and thus will suddenly shut off the inhibition from the amacrine cell to the bipolar cell and allow the bipolar cell to suddenly send a large amount of excitation to the ganglion cell's dendrite. This minimum amount of excita-

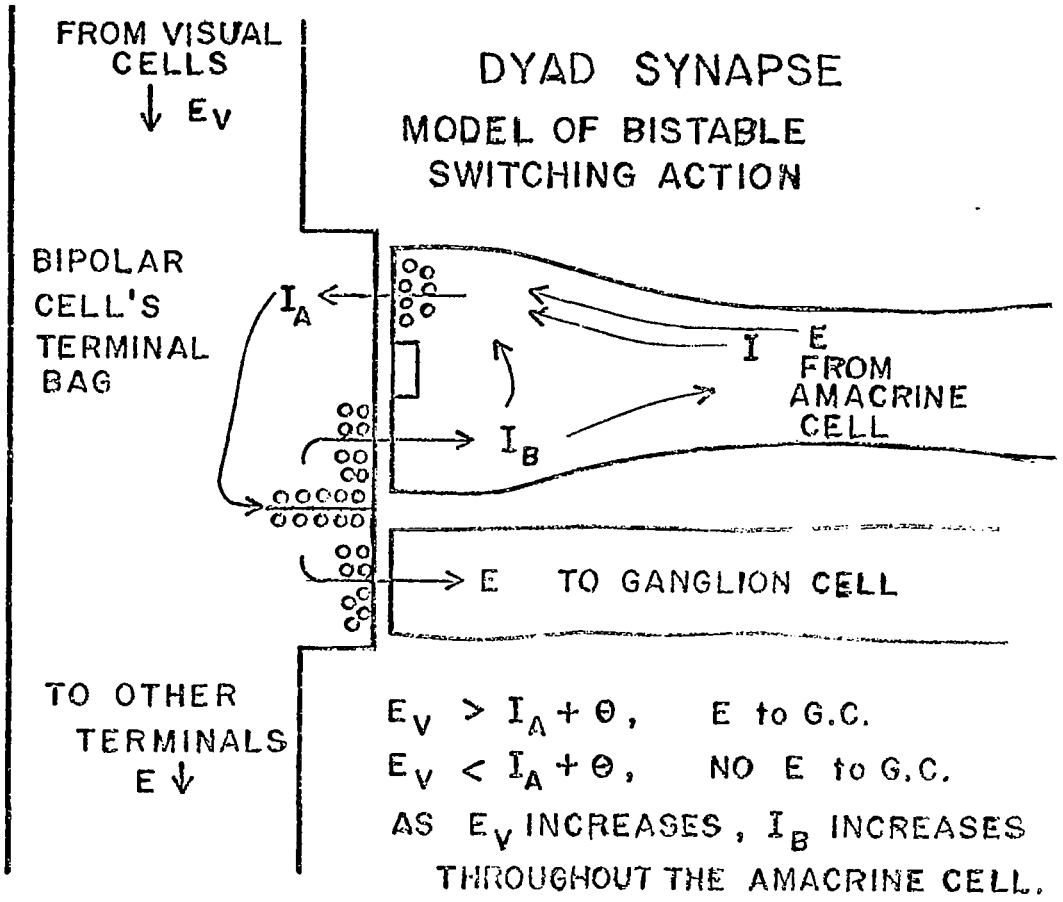


Figure 1

Diagram of so-called "dyad" synapse between a bipolar cell and nerve fibers from an amacrine cell and a ganglion cell in the primate's retina. Subscripts V, A, and B refer to visual cells, amacrine cell, and bipolar cell, respectively. E is excitation; I is inhibition; G.C. is ganglion cell. For explanation of the proposed gating mechanism of this synapse see the test. Θ is the depolarization necessary to start to activate the bipolar-to-ganglion-cell-and-amacrine-cell synapse.

tion sent to the ganglion cell might be of such a size that it will cause the firing of not less than several action potentials by the ganglion cell. In a similar way, as the excitation from the photoreceptors to the bipolar cell's terminal decreases, there will suddenly come a critical level of depolarization at which the inhibition from the amacrine cell can begin to be built up again and will suddenly turn off the bipolar cell's synapse and stop the sending of excitation to the ganglion cell. By changing the level of excitation within the amacrine cell, say, by inputs at distant branches of the amacrine cell, it would be possible to change the critical depolarization of the bipolar cell at which its synapse to the ganglion cell is turned on and off. This synaptic arrangement can thus be considered to act as a gating mechanism, and the amacrine cell is the device which sets the threshold for the gate. Such a mechanism would fit with the observation that the ganglion cell discharge often starts at a rather high frequency of discharge and stops suddenly with very little change in its frequency during the firing period.

(e) Lateral inhibition. - In many sensory systems, it has been found that activity in any given neuron tends to be inhibited upon stimulation of neighboring neurons. This phenomenon is termed lateral inhibition.

It has been found for the Limulus eye (82) that the threshold for firing of the first-order axon is always lower than the threshold to cause spread of inhibition from that axon to neighboring axons. From work on the frog's eye, one may believe this to be true of the vertebrate retina as well. It may be that this is the usual arrangement in lateral inhibition systems, but that has not yet been proven. This would fit the idea that inhibition acts only to modulate the excitation and does not carry information when alone. Notice that the touch receptor system described in a preceding section, which operated by converging excitatory influences alone, was designed to model and was tested for, only near-threshold conditions of stimulation. That is, it was tested for the conditions under which each first-order neuron fired only a single action potential upon stimulation. One suspects that if tested at larger intensities of stimulation, a lateral inhibition will be found to act in this system, also.

(1) Disinhibition. - In the Limulus lateral eye it has been shown that when additional first-order neurons are stimulated in the vicinity of an interacting pair of neurons, and are too far from one of them to affect it directly but near enough to the second to inhibit it, the frequency of discharge of the first increases as the inhibition exerted on it by the second decreases. This release from inhibition is called disinhibition. It can occur in inhibitory systems whether or not they involve lateral inhibition (83).

If the two inhibitions involved are of the divisive type, then disinhibition provides a mechanism for multiplication. For example, if activity X is inhibited by activity Y and this is disinhibited by the activity Z, then the final activity is kXZ/Y, where k is some constant. Thus, the product XZ is formed. The operation of multiplication must be performed in the calculation of correlations. It has been suggested (84) that such calculations are

commonly performed by the nervous system in the recognition of patterns, and that this is the reason for the widespread occurrence of disinhibition in the nervous system.

(2) Feedback type of lateral inhibition. - Disinhibition can occur in a lateral inhibition system only when the inhibitory influence that is exerted by a given order neuron depends on its own activity, which is the resultant of the excitatory stimulus to it and whatever inhibitory influences may, in turn, be exerted upon it by other same-order neurons. This mechanism of lateral inhibition may be termed "feedback lateral inhibition". See also (85).

(3) Feedforward type of lateral inhibition. - Another form of lateral inhibition is possible in which the output to the next order (next stage) neuron is reduced by inhibition from its neighboring same order neurons, but its own inhibitory output to its neighbors is not affected by this inhibition. Such a feedforward type of lateral inhibition will show a sharpening of spatial intensity gradient, but will not show any disinhibition.

In a study of the firing of the first order auditory neurons, which connect directly to the receptor hair cells of the cochlea, it has been shown (80) that the response of the fiber to a given audio frequency tone can be greatly reduced by the simultaneous presentation of a different frequency tone. Since each of these tones acts on hair cells at a different spatial location along the cochlea, this is a form of lateral inhibition. The researchers tested for disinhibition, but were unable to demonstrate it. It is possible that in this case the lateral inhibition is of the feedforward type.

(f) Categorization in the cat's visual system. - The cat's visual system has been studied in more detail than that of any other animal. The retinal ganglion cells have nearly circular receptive fields containing a central region surrounded by a peripheral zone. There are two classes of ganglion cells. One type will discharge at the onset of illumination if the illumination is confined to the central zone, and at the cessation of illumination if the illumination is confined to the peripheral zone. If the illumination reaches both central and peripheral zones, the result is a weak discharge which may occur at either the onset or cessation or both, depending upon the relative illuminations of the two zones. There is a mutual inhibition exerted by the excitation of each zone upon the effect of the other on the ganglion cell's discharge. The second class of ganglion cell is the same except that the zones for discharge at the onset and cessation of illumination are interchanged (86).

The ganglion cells connect to cells in the lateral geniculate body of the brain. The cells there have receptive fields like those of the ganglion cells, except for being a little smaller in diameter. These geniculate cells connect to cells in the cerebral cortex in what is called area 17 or the striate visual cortex (87).

The cortical cells in area 17 exhibit two types of receptive fields, the simple and the complex. The simple cortical cells have receptive fields made

up of two or three zones having opposite response types and separated from each other by straight line boundaries. Thus, in one zone of the receptive field, the simple cortical cell will give a response at the onset of illumination, called an "on" response. An adjacent zone gives a response only at the cessation of a light flash, this being called an "off" response. These simple cortical cells have the axes of their receptive fields in all possible directions on the retina. Diffuse illumination of the whole receptive field produces very little, if any, response at the cortical cell. The most effective stimuli of these cortical cells are bars lying on the axis of the receptive field with an edge at the boundary between zones; in the on zone either a light bar for an on response or a dark bar for an off response, and vice versa for the off zone.

The complex cortical cells of area 17 also respond best to bars or edges which are oriented in a particular direction for a given cell. However, for these cells it is not required that the edge lie along a particular line in the receptive field; it can be anywhere in the receptive field provided it has the proper orientation. It is as if each complex cortical cell has combined the outputs of a number of simple cortical cells all having the same orientation of their receptive fields. This idea is given support by the finding that the cells of the cerebral cortex in that region are arranged by function into columns extending perpendicularly through the thickness of the cortex, these columns being irregular in shape and approximately one millimeter in diameter. Within each column all the simple and complex cortical cells have the same orientation of their receptive fields (88).

The cortical cells of area 17 are connected and mapped in a mirror arrangement to cells in the neighboring area 18 of the cerebral cortex. These in turn are connected in a mirror arrangement to the cells in the neighboring area 19 of the cortex. The cortical cells of area 18 have receptive fields that have been termed "hypercomplex" (89). They will respond to the presence of two edges arranged either parallel or at right angles to each other and with a rather fixed separation and a specific orientation, although they can be placed that way anywhere in the receptive field. It would seem that such receptive fields could be built up by combining the outputs of two complex cells with receptive fields having either the same or perpendicular orientations and covering either the same or overlapping regions of the retina. This idea is borne out by the finding that adjacent columns in area 18 generally have at least one of their visual field axes in common and the other axis is often at right angles.

In area 19 are found both the hypercomplex cells and "higher order hypercomplex" cortical cells. These latter cells have receptive fields which are even larger in extent than those of the other cells, and give a maximum response to three edges suitably arranged with respect to each other anywhere in the receptive field. Usually, two of these edges are parallel to each other and the third at right angles. The hypercomplex cells found in the same column as the higher order hypercomplex cells have the same orientation of receptive field axis, and it is probably by the combination of the outputs of several of them that the receptive field of the higher order hypercomplex cortical cell is constructed. At these cells, there already exist

such abstract concepts as U shape, or T shape, or corner and side, which can be varied greatly in location on the retina and somewhat in size and still give an optimal response. In area 19 there are again columns of cortical cells which represent the different orientations possible in the receptive field, and these columns are so arranged across the cortex as to map, with great overlap, adjacent regions onto the cortex.

The obvious extrapolation from these findings is that by further combinations of these categories of information it would be possible to build up still more abstract categories such as "square" or "table", etc. However, many steps of generalization would be needed to get from the U shapes of area 19 to the concept of a specific familiar table. From area 19 there are only a very few cortical steps involved at most (associative cortex and sensorimotor cortex) before signals for action responses are sent on to the motor nuclei. It, therefore, seems unlikely that the additional categorization is done by the "one step per cortical cell" mechanism described by Hubel and Wiesel. This is an important question that needs much further investigation.

A possible suggestion is that two further mechanisms are involved in the categorization of sensory inputs by cerebral cortex. The first mechanism is lateral modulation. It is known that the cortical input is received by the dendrites of the pyramidal cell and the output is via the axon of the same cell (94). Thus, all calculations of categorization must be done at the pyramidal cell. The usual interpretation of Hubel and Wiesel's work is that a given pyramidal cell can perform only one particular categorization (as determined by its geometry and the types of simpler categories signalled by the sensory fibers synapsing with it). One can suggest that the categorization performed by that given pyramidal cell can be further modified by the other inputs to the pyramidal cell. These inputs include (a) inhibition by interneurons which are excited by other pyramidal cells as much as 10 mm distant along the same gyrus (90, 91, 92, 93) and in symmetrical areas of the contralateral cortex (90, 98). This could mean that the particular light pattern which would most excite a pyramidal cell in area 18 would be modified by the pattern of stimulation reaching the other pyramidal cells in <u>both</u> areas 18. The inputs from the specific thalamic nucleus (e.g., the lateral geniculate body in the case of the visual system) terminate at synapses on the vertically oriented apical dendrites of the pyramidal cell (97). Other inputs (b) from nonspecific thalamic nuclei terminate on the apical dendrites (69, p. 428). These may include inputs from a number of sensory systems (95). Thus, the activity of other sensory systems could modify the pyramidal cell's categorization. There is also input from the reticular formation (discussed later) to the apical dendrites of the pyramidal cells. It appears that at a low level of reticular formation activity there is a subliminal excitability increase in the pyramidal cell. At higher levels of reticular formation activity, there is a reduced cortical activity (69, p. 437), possibly as a result of the intracortical inhibitory interneurons.

There is a second mechanism which can be suggested as a means of modifying the pyramidal cell's activity so that it responds to a more complex and specific category of sensory input. This is a recycling of the pyramidal cells' outputs back through the cells (that is, iterative processing) with

additional modification occurring during each cycle. The evidence for such a mechanism is weak, but it is known that an isolated piece of cortex will produce rhythmic bursts of discharges upon receiving a stimulation (69, p. 425). This also happens in the intact cortex following a single shock to the connecting nonspecific thalamic nuclei (96), and to some extent at the visual cortex following a light flash to the retina (99). Such a mechanism might explain the comparatively long time required to perceive a complex stimulus as belonging to a specific complex category.

(g) Habituation. - After the nervous system has categorized the incoming information from the environment, it is necessary next to decide whether this represents new conditions which require new actions or is simply a continuation of a previous condition and can be met by a continuation of the previously decided upon actions. For example, a hoop rolling past will be seen first as an ellipse, momentarily as a circle, and then again as an ellipse, and as varying sizes and orientations. Presumably, the processes of categorization of the form described previously will simply be signaling "circle moving past at constant velocity".

There exist in the nervous system various mechanisms for determining whether the information signaled by a neuron is not really new information and therefore is a redundant signal. Habituation of a neuron is the reduction in its response to a given stimulus upon repetition of that stimulus; and provides a mechanism for reducing redundancy in the signals.

A "sameness detector" cell has been found in the optic tectum of the frog (100). It is silent until an object is brought into the receptive field and then the cell fires a burst of impulses whenever the object shows angular or linear acceleration of its movement. The cell continues a steady muttering of impulses as long as the object remains in the visual field. If a second object enters the visual field, the cell might latch onto the second object and then would ignore changes in motion of the first object and only report changes in motion of the second object. If the second object was held still for awhile, and the first object moved, the cell would then latch back onto the first object. It would seem that the function of the cell is to keep track of whatever is the most actively changing object in the visual field.

A "newness detector" cell also has been found in the optic tectum of the frog. It has a receptive field about 30° in diameter. If an object is moved along a particular diameter of the receptive field in one direction, it evokes a response; but, repeating that movement over the same path in less than five to ten seconds brings no response. Even with a 20 second delay to the second stimulation, the second response is much reduced below the first. However, if the object is moved at 90° with respect to the first path, there is a full response. Both paths can be adapted within one second, whereupon a third movement along any direction will not be effective for the next ten seconds.

7. Evaluation of Input Information - 'Subgoals' as Hierarchical Entities

After having categorized the input information from the environment and evaluated its novelty, the nervous system must next decide the relevance of this information for the attainment of the subgoals having the highest priority at that moment. In this evaluation, information as to the internal state of the organism must also be considered.

(a) Assignment of priorities to subgoals. - At any instant the organism has one or more tasks which it is trying to perform. One may refer to such tasks as subgoals, since they are only part of the behavior involved in the overall goal of survival of the species (i.e., they exist within a hierarchical structure). Each subgoal can have different priorities with respect to other subgoals and each can have certain periods within which that subgoal must be achieved. For example, a hungry animal might have several days in which to locate food and ingest it. On the other hand, a person in the path of an approaching car may have only a fraction of a second in which to meet the subgoal of getting out of its way. Obviously, a hungry person in the path of the car will assign a higher priority to avoiding the car than to getting food, and will not stop to pick up the piece of food lying on the road in front of the approaching car.

Note that several types of sensory information are used in establishing these subgoals and their priorities. The subgoal of hunger might depend upon stimulation of pain receptors in the stomach, upon receptors in the brain which are stimulated by a low level of glucose in the blood and upon activation of hormonal mechanisms. The subgoal of avoiding an approaching car depends upon visual and auditory inputs which must go through a great deal of categorization for the organism to be able to recognize that that particular situation is a dangerous one which must be avoided very quickly. Such a subgoal is a rather shortlived one, and depends entirely upon information carried by the nervous system. A subgoal like hunger is a longer lived one, and depends both on information carried by the nervous system and upon information carried by the hormonal system.

(b) The reticular formation of the vertebrate brain. - The question then is what organizational structure in the brain receives all these different types of inputs and can correlate them to establish the subgoal. The reticular formation meets this criterion. It consists of a central core of neurons lying in the spinal cord, brainstem and the midbrain. The lower two thirds of the reticular formation, called the brainstem reticular formation, receives inputs from peripheral receptors. These sensory inputs have remained little-categorized. It also receives little-categorized inputs from the visceral receptors.

The upper third of the reticular formation in the head is called the thalamic reticular formation. It receives inputs from branches of the nerve fibers that supply categorized information (from sensory systems served by the peripheral receptors; at the skin, muscles, and joints) to the so-called specific nuclei in the thalamus. These nuclei further categorize the informa-

tion and send it on to the cerebral cortex for still further categorization. It also receives inputs from the brainstem reticular formation (101).

(1) Thalamic reticular formation. - The reticular formation of the thalamus differs in organizational pattern and functional role from the remainder of the reticular formation. It is partially isolated from the reticular structures of the lower two-thirds of the brainstem. It, much like the lower reticular formation, consists of a group of dendritic fields stacked like a pile of chips along the long axis. Ascending axons from the posterior two-thirds of the brainstem run roughly parallel to this system and provide one of its most important sources of inputs. It also receives many inputs of branches of axons from the fiber tracts that carry categorized sensory information toward the cerebral cortex. The reticular formation receives directly signals from the sensorimotor and adjacent portions of the cerebral cortex. It also receives signals from the cerebral cortex via the nuclei of the limbic system. There are also input connections from the cerebellum.

The reticular formation has outputs to adjacent nuclei such as the ventral-anterior nucleus reticularis thalami. The latter nucleus is a shell of neurons surrounding the thalamic reticular formation on its upper three sides so that all the connections from the thalamus must go through this nucleus. It is, accordingly, in a critical position to monitor and modulate most thalamus-to-cortex and cortex-to-thalamus interactions. Most of its outputs are distributed to the thalamus and to the mesencephalic tegmentum, providing feedback loops to the entire reticular formation.

(2) Reticular formation of the brainstem. - The dendrites of reticular core cells appear to be relatively straight, long, and unbranched. Such dendrites spread little along the long axis of the brainstem, but spread out greatly across the transverse axis of the brainstem. The appearance of the dendritic trees has been likened to a stack of poker chips. Thus, each particular reticular formation cell looks at only a limited series of inputs from a single level along the input continuum of the brainstem's long axis. Each of the particular neurons can be thought of as receiving its own particular sampling of the many different inputs to the reticular formation, and of acting as an integrating subcenter upon this particular combination.

The reticular formation of the brainstem may be thought of as overseeing all incoming and outgoing information-carrying systems. The evolutionarily older systems that carry less well-categorized information are more thoroughly represented than the newer, more categorized, information input systems. This suggests that information with a high degree of locus and categorization is not crucial to the operation of the brainstem reticular formation.

(3) Reticular control of sensory input. - The reticular formation of the brainstem is able to block by inhibition the categorized sensory inputs to the brain (69, 102, 103, 104). It will do this for a given type of sensory input if that input's signal is simply a repetition of previous identical signals (habituation). It will also do this if the signals arriving from other types of sensory inputs are judged to require more attention by the

organism. The usual example of this is that the evoked potentials (massed discharge of input neurons) measured in the lateral column of the spinal cord in response to tactual stimulation, and also the discharge of the cochlear nerve in response to sound, are both diminished in a cat when the cat is attentively sniffing a fish odor. The mechanism of this inhibition has been shown in some cases to be postsynaptic inhibition at the first synapse (between receptor and neuron) of the sensory pathway. There is also evidence that this inhibition can act at the second synapse of the sensory input chain. The latter inhibition is generally produced by activation of the thalamic reticular formation.

Cortical evoked potentials are more prominent during deep anaesthesia than during light anaesthesia; there is a marked facilitation induced by barbiturate anaesthetics. Simultaneously the sensory signals along the incoming categorized sensory pathways are enhanced. The conclusion is that barbiturate anaesthesia reduces the inhibitory influences which normally act during wakefulness upon the incoming sensory pathways. Habituation is found to be released and prevented by these anaesthetics.

If the excitation from the thalamic reticular formation to the cerebral cortex is blocked, the animal tends to go to sleep. This can occur by inhibition of the thalamic reticular formation via activity of the brainstem reticular formation. It can also occur as a result of barbiturates which act directly on the cells of thalamic reticular formation. It can be concluded that excitation from the thalamic reticular formation to the cerebral cortex is necessary to maintain a wakeful state in which the cortex will respond to the incoming sensory signals. Part of this excitation seems to result from recognition of the novelty or sudden change of the sensory input. This is signalled by a large increase in cortical sensitivity to that particular input, this increase being called the orienting reflex. Habituation occurs for those sensory inputs which are repetitions of the previous inputs.

While habituation to categorized sensory inputs is controlled by the thalamic reticular formation, there is much evidence that the brainstem reticular formation is also capable of causing habituation of various sensory inputs. Most habituation is lost if the latter structure is damaged. It has been found that lesions at each level in the brainstem reticular formation have more effect (reduced inhibition and habituation) on a particular type of sensory input than on the other types of sensory input. This is thought to result from the neurons in each region of the brainstem connecting predominantly with one type of sensory input.

Habituation can also occur at lower levels in the reticular formation. For example, habituation to tactile stimuli has been observed in the isolated spinal cord. The habituation controlled by the brainstem reticular formation acts in addition to or via such a local habituation mechanism.

As would be expected, selective facilitation of one sensory input is often accompanied by the selective inhibition of other types of sensory inputs. For example, electrical stimulation of the brainstem reticular formation inhibits the activity of lateral geniculate neurons (105).

(4) Rapidity of sensory input to the reticular formation. - The various spinal tracts which carry proprioceptor, cutaneous receptor, and joint receptor information to the cerebellum also send such information to the reticular formation, and have a continuous background activity of 25-30 impulses per second (106). The vagus nerve also sends visceral sensory information to the midbrain and the reticular formation, and has a continuous and inherent rate of impulse activity of very close to 28 per second (107). Such rapid rates of continuous activity allow, by modulation of that rate or of the summed amplitude of the action potentials, a means of rapidly informing the reticular formation of any increases or decreases in sensory activity throughout the body.

(5) Hormonal modulation of reticular formation activity. - It has been found that the threshold for activation of the reticular formation drops precipitously at full estrus and becomes two to three times higher during the non-estrus days of the animal's sexual cycle. On the basis of this and other evidence, it has been proposed that the hormones and the sensory input both act upon the reticular formation to maintain a state that leads to the initiation of instinctual behavior (103).

(6) Cyclic phenomena. - In about 80 per cent of the reticular neurons whose electrical activity was tested, there was a succession of periods during which the cells appeared alternately sensitive to sensory bombardment from the external milieu (mild shocks to a leg nerve, the sciatic) and from the internal milieu, as exemplified by rhythms which followed the respiratory cycle (101). During these respiration-sensitive periods, the reticular neurons were completely insensitive to signals from the peripheral receptors. Similarly, during the sciatic-shock-sensitive periods, no traces could be seen of the effects of visceral activity. Each neuron seemed to follow a unique temporal pattern with the periods ranging in general from one half to three hours. There was no discernible relation of these periods to sleep-wakefulness cycles nor were there indications of general physiological changes accompanying the changing of sensitivity.

One suggestion to explain these cyclic phenomena has been that the presynaptic inputs to the neurons are being manipulated by a population of pacemaker neurons that periodically change the sensitivities at the synapses from the different sensory sources. Another suggestion has been that the input sensitivities of the reticular neurons are being changed by interaction with the oligoglia cells (connective tissue cells found in the nervous system) which surround them, and which have been shown to receive branches from the same input neurons that connect to the reticular neuron. A third suggestion is that those local arrays of reticular cells whose information input is of a more biologically urgent nature than that of the adjacent grouping of cells take over control of the sensitivity of those neighboring cells to the various types of input signals (101). (We can add a fourth suggestion which is that the cyclic activity represents a mechanism for selecting from the varieties of sensory inputs only those which are relevant to the subgoals of the organism as of that instant. The time periods over which an organism sticks to a given subgoal seem to tie in with the visceral states of the organism and fit the time periods of the various hormonal control systems. This would

be consistent with the periods of one half to three hours observed in the reticular neurons. Therefore, one can suggest that the hormonal system may act, possibly via the blood stream, to change the relative effectiveness of the synapses on the reticular neurons from the different types of sensory inputs. A check on this notion would be to see whether reticular neurons which were fairly well separated from each other in the brainstem showed simultaneous alterations of their sensitivity to particular types of sensory inputs.)

8. Output Systems

The central nervous system has three main branches; the sympathetic and parasympathetic nervous systems, which together make up the autonomic nervous system, and the somatic nervous system. The somatic system is primarily concerned with the sensory inputs from the receptors in the skin, the muscles and joints and with the outputs through the muscles. The autonomic nervous system sends outputs to and controls (a) those muscles not concerned with posture, and (b) the secretory activity of a number of glands.

The autonomic systems are usually spoken of as having no sensory function, but this notion is being increasingly challenged, particularly by work of the East European neurophysiologists. It has been found that the extraspinal autonomic nervous system can transmit information about visceral conditions in the pelvic region on to the level of the head (108, p. 207). There is also evidence in the lower vertebrates that during evolution as the size of the cerebellum (the center in the brain that coordinates all muscular actions) increased, so did its connections with the trigeminal and vagal nuclei and the hypothalamus. Since the size of the cerebellum must be related to the amount of sensory input, this suggests that the trigeminal and vagal nerves of the autonomic system are supplying sensory inputs to the cerebellum (108, p. 290). More recent work (107) has shown that there is increased activity of the vagal nerve, even when its connection with the brain is cut, upon distention of an animal's stomach by inflating a balloon within it. The vagal nerve seems to connect with receptors in a great many of the visceral tissues and it seems to send this information to a number of midbrain structures, including the reticular formation.

(a) Skeletal muscle control system. - The major output by which a vertebrate animal can affect the relations between it and its environment is movement. This movement is brought about by activation of skeletal muscles (the muscles acting on part of the skeleton).

The skeletal muscles are assemblies of muscle cells, each connected to at least one nerve fiber, whose activation evokes a contraction of the muscle cell. The nerve fibers are the axons of (alpha type) motoneurons, whose cell bodies are located in the spinal cord. The mechanical state of each muscle is monitored by several types of sensory organs (called proprioceptors). They detect such quantities as the angles of the bones forming a joint (by the Ruffini endings in the joint), the tension in the muscle's attachment to

a bone (by the Golgi tendon organs), and the stretch within the muscle (by the muscle spindle).

Axons from the proprioceptors terminate either directly or via interneurons onto the motoneurons in the spinal cord. The activity of the motoneurons can be modified by activity of the proprioceptors. Change in position or stretch of a muscle can change the activity of its proprioceptors and via them change the activity of motoneurons and of the muscles to which they connect. Such simple interactions are termed "reflexes". For example, application of tension to a muscle can by the stretch reflex increase the activation of the motoneuron to that muscle and increase the tendency of the muscle to contract. In this way the muscle opposes the increased pull and holds its joint fixed in position. This reflex plays a major part in maintaining body posture against the pull of gravity. For further details of such reflex circuits, see reference (135).

The output of a muscle spindle for a given muscle stretch can be changed by changing (say, by a neuronal connection from a brain center) the activity of the (gamma type) motoneuron that controls the contraction of muscle cell within the muscle spindle. Changing the input to the gamma motoneurons of a muscle changes the muscle's equilibrium position in the stretch reflex. The muscle can thus be used as a remote-controlled servomechanism, with control exerted by other reflexes in the spinal cord or by the brain centers.

The brain centers coordinate muscle activities on the basis of sensory inputs, muscle conditions, body position, and desired subgoal. The sensory inputs include those from categorized sensory inputs and their further categorizations by the cerebral cortex, as well as uncategorized inputs. The muscle conditions are signalled by branch axons from the proprioceptors to the brain. Body position is signalled by proprioceptors of the joints and by sensors of: the direction of gravitational pull (otolith organ) of inner ear, linear acceleration (otolith organ), and angular accelerations (semicircular canals of inner ear). These latter sensors connect to the vestibular nucleus of the brain.

Voluntary actions are probably initiated in the frontal cerebral cortex or other association areas of the cortex, sent to the sensorimotor cortex, and there relayed in general terms of the desired result to the basal ganglia. These ganglia are the headmost centers of an evolutionarily older motor control system. The basal ganglia reinterpret the general terms into the detailed command routines formulated by the lower centers (brainstem and spinal cord). The reticular formation, cerebellum and vestibular nuclei work as a closely correlated group in the routine maintenance of equilibrium and walking. The cerebellum monitors the voluntary muscle movements and modifies them to attain the desired results, signalled it by the cerebrum. More detailed descriptions of the skeletal muscle control system are given in reference (109).

(b) Central control of the autonomic system. - The hypothalamus has been found to be the part of the brain which exerts the greatest control over

the activities of the autonomic nervous system. Stimulation of various parts of the hypothalamus results in activation of the autonomic system. Stimulation of the ventral hypothalamus results in activation of the sympathetic system. Stimulation of the anterior hypothalamus (more likely of the cells just above (rostral) to it) causes activation of the parasympathetic system. Stimulation of other parts of the brain, such as the cerebral cortex, can act, usually via the hypothalamic region, to cause activation of the autonomic nervous system.

(1) Sympathetic control of sensory inputs. - It has been shown that the electrodermal response (also called the psychogalvanic skin response and the galvanic skin response, GSR; a decrease in electrical resistance of the skin) in humans is closely correlated with the degree of behavioral arousal. Both the GSR and arousal are increased by stimulation of the sympathetic nervous system. Stimulation of the sympathetic output nerves has been observed to enhance the response of such receptors as, touch, chemical, muscle spindle, and gustatory. (See also reference (110)). All of these findings are consistent with the idea that sympathetic system enhancement of peripheral sensory input might play an active role in behavioral arousal of vertebrate animals.

(2) Sympathetic system's outputs. - Each of the sympathetic output fibers connects with effector cells (cells that carry out some action, such as movement or secretion) over fairly wide regions in the body. Cannon (111) gave the general rule that activation of the sympathetic system prepares the body for increased activity under emergency conditions. It does this by reducing all activities that do not contribute to muscular work; for example, by constricting the blood vessels to the viscera and thus reducing their blood supply, by inhibiting movement of the gastrointestinal tract (and thus reducing work done in digestion) and by inhibiting most gastrointestinal glandular secretion, thus conserving water for the sweating necessary to keep the body temperature down during muscular exertion. It facilitates all activities that aid muscular effort; for example, by increasing heart rate, by dilating the blood vessels to the skeletal muscles so as to increase their blood supply, by activating the sweat glands so as to keep the body cool, by dilating the bronchioles so as to increase the air ventilation to the lungs, by contracting the spleen so as to force more blood into circulation, and by causing secretion of epinephrine from the medulla of the adrenal gland. The cells of the adrenal medulla are specialized nerve cells, much like those which make the output connections of the sympathetic system. The adrenal medulla's cells secrete epinephrine and related compounds, much as the regular nerve cells do at their terminal synapses. These compounds are released into the blood circulation and have a number of effects. One of the effects is to act on the synapses that the regular sympathetic output neurons activate, and prolong the activation of these effector cells. The secretions also cause the release into the blood stream of increased levels of free fatty acids and glucose, and act on certain enzyme mechanisms to increase oxygen consumption and energy production in the body.

The sympathetic system also has continuous, non-emergency discharges to many other structures, such as the heart and the pupil, causing increased

heart rate, and dilation of the pupil, and assisting in their regulation. Thus, it acts as a general arousal system which prepares the body for rapid, large responses.

(3) Parasympathetic system's outputs. - The parasympathetic system has its nerve fibers terminating on effector cells within rather limited regions. The actions of the parasympathetic system seem to be more localized than those of the sympathetic system. Its output synapses use acetylcholine as the chemical transmitter. Its effects are often the reverse of those of the sympathetic system. Its activity causes increased motility in the gastrointestinal tract, increased secretion by the digestive glands, a slowing of heart rate, and a constriction of the pupil. All these effects are controlled by the cranial portion of the parasympathetic system. The sacral portion (near the lower end of the spine) is largely concerned with the elimination of material from the body. Its activity causes contraction of the bladder and inhibition of the urethral sphincter to allow the voiding of unwanted fluid. It causes increased tone and motility of the colon, so as to reject unwanted solid waste. The parasympathetic system has no connections to the blood vessels of skeletal muscles.

Thus, the parasympathetic system prepares the body, not for instant strenuous action, but for getting on with the other activities utilized in survival, such as digestion, food-hunting. In a way, it counteracts the sympathetic system and allows the body to switch to non-strenuous activities.

(c) Maintenance monitoring. - For continued operation of the nervous system, the physical and chemical conditions within it must be maintained within those limits which allow the neurons to function. Therefore, one would expect the nervous system to contain receptors to monitor those conditions and connecting outputs to trigger needed corrections. For example, one would expect monitoring of temperature, pressure, glucose concentration, oxygen concentration, and carbon dioxide concentration.

(1) Maintenance monitoring by the hypothalamus. - The hypothalamic nuclei have been shown to contain receptors which are sensitive to chemical and physical conditions in the brain. For example, the feeding-satiety center of the ventro-lateral nucleus has been shown to have receptors which are activated by glucose in the blood passing through that nucleus. (It is our suspicion that this is not part of the regular control mechanism of satiety vs. feeding, but instead is one of the maintenance monitoring mechanisms by which the nervous system ensures that the brain is supplied with the physical and chemical conditions that enable it to function. Any time the monitors indicate that one of those conditions is approaching a danger level, the receptors in the hypothalamic nuclei activate those nuclei to act as override controls on the normal regulatory systems so as to bring that condition back within safe limits. Thus, we suspect that the blood glucose detectors of the hypothalamic feeding-satiety center are activated only when the blood glucose level of the brain begins to approach a dangerously low level.)

The hypothalamus contains many other maintenance monitoring systems. It seems to contain temperature-sensitive receptors, probably in the lateral por-

tions of the caudal nuclei of the hypothalamus, which monitor the temperature of the brain. They signal the need for increased heat, which is accomplished by constriction of the surface blood vessels, by erection of the surface hairs to provide increased thermal insulation, by shivering of the skeletal muscles, and by an increased metabolic rate of the cells through the release of thyroxin hormone. Nuclei in the anterior hypothalamus seem to be responsible for getting rid of excess heat; this is accomplished by dilation of the surface blood vessels, and by sweating, to increase cooling by evaporation.

It has been suggested that certain structures in the hypothalamus are sensitive to changes in the osmotic pressure of the blood. These signal the water concentration in the blood, and can act to increase this water concentration by stimulating drinking behavior and can, by acting via the pituitary, cause release of a hormone which causes the kidneys to reduce their water loss to the urine.

(2) Maintenance monitoring by other parts of the brain. - In addition to those found in the hypothalamus, other important monitors of the brain's physical and chemical conditions are found in the medulla of the brainstem. There are receptors in the medulla which are activated by an increased concentration of carbon dioxide in the blood; they, in turn, activate motor centers to cause increased respiration. Receptors in the carotid artery (one of the main blood supplies to the brain) signal changes in blood pressure to a control center in the medulla, which then adjusts the amplitude of heart contractions and the dilation of the blood vessels throughout the body to restore the blood pressure to the appropriate level (69, pp. 527-548).

9. Coordination of the Neural and Hormonal Systems

(a) Neuroendocrine integration by the hypothalamus. - If one considers the extent to which both the neuronal and the hormonal control systems regulate body functions, it is not surprising to find that there are large parts of both systems which act as an interface between the two to coordinate their activities. An excellent introductory review of this interface is provided by Ganong (112), and more elaborate descriptions can be found in books on the subject (113). The major interface is between the hypothalamic region of the brain and pituitary region (hypophysis) of the hormonal system, which is located just under the hypothalamic region.

(1) Direct neuronal control of the posterior pituitary. - There are two groupings of cell bodies, the paraventricular nucleus and the supraoptic nucleus, in the hypothalamic region which have axons that run through, and end at, the posterior portion of the pituitary gland. When these neurons are stimulated, they release certain peptides from their terminals in the posterior pituitary directly into the capillary blood flow through the pituitary. These peptides include oxytocin and vasopressin. Oxytocin is known to affect the length of the estrous cycle of certain mammals, induce labor in pregnant mammals, and induce the ejection of milk from the mammary glands. It has been shown by lesions in different parts of the hypothalamic region that stimulation of both the supraoptic and paraventricular nuclei evokes no

ejection. Normal stimuli for that neuronal activity are such things as suckling by the young, of the mother. In some species distention of the vagina is effective. It is thought that mechano-receptors at these body surfaces are stimulated and conduct activity, possibly along a lemniscus pathway. Their exact path to the hypothalamic region is not known, though it is known that at many points in the limbic circuit (explained later) stimulation will cause the release of oxytocin.

Vasopressin, which is also called antidiuretic hormone, is produced by neurons in both the paraventricular and supraoptic nuclei. Its primary effect is to cause resorption of water by the kidney so as to reduce the amount of water in the urine. A typical response develops slowly over the course of a few minutes and persists ten to thirty minutes. Larger amounts of vasopressin cause a rise in the arterial blood pressure. The normal stimuli for the release of vasopressin are any changes in the body condition which make the retention of water an aid to survival. For example, vasopressin is released in animals which are deprived of water, or upon an increase in the osmolality of the blood plasma (probably detected by receptors in the carotid arteries), by hemorrhage to the point that the arterial blood pressure is reduced (probably detected by receptors in the carotid arteries), by hemorrhage to the point that the arterial blood pressure is reduced (probably acting via receptors in the left atria of the heart). Increased blood pressure is known to produce inhibition of vasopressin release. Painful stimuli can cause vasopressin release in the absence of other stimuli. Stimulation of the vagus is thought to cause vasopressin release. Thus, there are stimulating and inhibiting pathways via somatic and visceral nerves through pathways in the medulla and the pons (both being regions of the brainstem) and some points in the lateral reticular formation. A number of these pathways go to the midbrain and then are relayed to the hypothalamus, where stimulation of the paraventricular and supraoptic nuclei causes vasopressin release. The osmolality receptors which test the blood flow coming via the carotid arteries are thought to lie within the following parts of the brain: medial and lateral preoptic region, anterior hypothalamic region, paraventricular and supraoptic nuclei, central medial nuclei and the dorsal and lateral hypothalamic regions. Other portions of the brain are also stimulated by osmotic changes, but need not be involved in vasopressin release.

(2) Indirect neuronal control of the anterior pituitary. - The median eminence is that portion of the brain which lies between the pituitary gland and the hypothalamus. Many blood vessels, the portal-hypophyseal vessels, pass through the median eminence and divide into numerous capillaries. Nerve cells from the hypothalamus, and possibly from other portions of the brain, terminate their axons on the capillaries in this region. The capillaries, in turn, drain into the blood vessels (the portal vessels) which pass into the anterior lobe of the pituitary, and there divide into many capillary loops which in turn drain into the general blood stream. The evidence by now is rather convincing that the nerve cells secrete certain chemical agents into the capillaries of the median eminence, and these agents pass along the portal vessels to the anterior pituitary where they act on certain secretory cells of the pituitary. It has been suggested that the anatomical arrangements of the blood vessels are such that the chemical agents from cer-

tain neurons are carried only to particular groups of cells in the pituitary. Six different chemical agents, or factors as they are called, have been extracted to date. The first five act to cause the release of particular hormones by the anterior pituitary. The sixth acts to inhibit the release of a particular hormone by the anterior pituitary. These hormones are released into the general blood stream and circulated until they reach the body cells on which they act.

The six chemical agents (often called factors) are, first, the corticotropin-releasing factor. This causes the release from the pituitary of the adrenocortico-tropic hormone (ACTH). ACTH causes the outer section of the adrenal gland to release into the blood stream hormones which are primarily concerned with the regulation of cellular metabolism and resistance to stress. ACTH is also necessary for the initiation of milk production in mammals, and for the normal ability of animals to learn avoidance responses. The second neuronal factor is the thyrotropin-releasing factor. It causes the release from the pituitary of the thyroid-stimulating-hormone (TSH). TSH causes the thyroid gland to release a hormone, thyroxine. Thyroxine increases the oxygen consumption, increases metabolic rates, and increases the release of stored glycogen as free glucose in the blood. Thyroxine acts on the hypothalamus to reduce the thyrotropin-releasing factor. Thyroxine also acts on the anterior pituitary to reduce the release of TSH. Thus, there is a pair of feedback loops which tend to keep the level of thyroxine in the blood at a constant value.

The third factor is the growth-hormone-releasing factor which causes the pituitary to release growth hormone, also called somatotrophin (STH). STH is involved in the general growth of all cells and in regeneration of damaged tissues. It also causes cells in the islets of the pancreas to release a hormone, glucagon, which increases the rate of release into the blood of stored glucose. The fourth factor is the follicle-stimulating hormone factor which causes the anterior pituitary to release the follicle-stimulating hormone (FSH). FSH acts to cause growth and release of ova by the ovary in the female, and maintenance of the formation of sperm cells in the male. The fifth neuronal factor is luteinizing-hormone-releasing factor which causes the pituitary to release the luteal-stimulating hormone (LSH). LSH acts on the corpora lutea glands in the ovary to cause them to grow and to secrete progesterone, one of the sex cycle controlling hormones.

The sixth neuronal factor is the prolactin-inhibitory factor which causes the anterior pituitary to reduce the amount of prolactin hormone that it releases into the blood. Prolactin induces the secretion of milk in the mammary glands.

(3) Control via the intermediate lobe of the pituitary. - It is known that in those amphibians and reptiles which are capable of darkening their skin when placed on a dark background, the hormone controlling the skin darkening is released from the intermediate lobe of the pituitary gland. The intermediate lobe's activity seems to be controlled by light stimulation of the retina and by a hormone which is released by the pineal gland and which inhibits release by the intermediate lobe of its hormone. The pineal gland

is located at the base of the brain, but it has been shown in some species of animals to have photoreceptor cells much like those of the retina and to receive light through the skulls and brains of the animals.

There is incomplete evidence that the pineal gland, particularly in birds, signals to the animal, probably via its control over all lobes of the pituitary gland, the arrival of certain seasons of the year calling for certain activities. For example, it signals mating season for birds, and probably also migratory seasons. A probable basis of this signalling is the ratio of the lengths of night to day of the various seasons of the year as detected by the light receptors of the pineal gland.

(b) Direct neuronal control of hormone production. - The autonomic nervous system connects via the sympathetic splanchnic nerves to the medullary cells of the adrenal gland. Activation of those nerves causes the release by that gland of epinephrine and norepinephrine (also called adrenalin and noradrenalin) into the blood stream. The splanchnic nerves also control the release of gastro-intestinal hormones by cells in the gut, and the release of insulin by the beta cells in the islets of Langerhans in the pancreas.

(c) Feedback actions of hormones on the brain. - As has been mentioned above, thyroxin can inhibit the release of the thyroid-hormone-stimulating factor by the hypothalamus, as well as directly inhibit the secretion of the thyroid-stimulating hormone by the anterior pituitary. It is also known that certain gonadal hormones can inhibit the hypothalamus to cause both a decrease in its release of gonadal gland stimulating agents by the anterior pituitary and also a change in its nervous control of sexual behavior. This feedback mechanism seems to be an essential part of the cycles of sexual behavior in animals.

(d) Neuronal and hormonal control of developmental processes. - The interacting neuronal and hormonal mechanisms play a major part in the patterning and timing of developmental processes. For example, damage to the hypothalamus can cause precocious puberty. It has been found by experiments on animals that if the brain of a female animal is exposed to male hormone (e.g., testosterone) at a critical time in its development, it does not develop as a female brain, but instead becomes a male brain. The animal having that brain will not show any female sexual behavior or any of the cyclical sexual behavior associated with females. Furthermore, if the pituitary from a newborn animal is transplanted into such a brain, it will release only those hormones and only in that pattern associated with the male. From this, it can be concluded that the brain is made to be either a male or female brain by its early exposure or lack of exposure to male hormones, and that it is the brain that tells the pituitary to put out either the male or female pattern of hormonal control. In addition, the brain neuronally controls the pattern of sexual behavior. The tie-in of the neuro-hormonal controls with the genetic control system is just being uncovered. Recent work has demonstrated in insects that hormones can stimulate the synthesis of messenger RNA. This provides the mechanism by which the hormone can cause the cell on which it acts to change its content of enzymes, and thus to change its metabolic rates, or to produce new chemical products, or to grow into a different

shape and structure, etc. Such a mechanism probably exists in the vertebrates as well, and provides the final interface between the neuronal, hormonal, and genetic systems of control.

10. Selection of Action Routines for Attaining a Subgoal

(a) Evidence for action routines. - By action routine is meant a combination and sequence of actions that accomplish a task. Thus, lifting a cup to the mouth is a simple action routine, drinking from a cup is a more complex one, and eating a meal is a combination of many such action routines. A "posture" is an action of the body on the body (114); a "motor response" is usually taken to be an action of the body on the environment.

(1) Direct control of the brain. - It has been found that electric stimulation of certain parts of the brain will lead to certain patterns of muscular activity. For example, electric stimulation at any of many points caused a monkey to produce a sound. Practically every type of sound the monkey is known to make was elicited. Often specific vocalizations were elicited by stimulation in definite loci in the brain (115).

The earliest work was done by W. R. Hess (for a review of his work, see reference (116)). He demonstrated that electric stimulation of specific cerebral structures, such as the hypothalamus or midbrain nuclei, would influence body posture, equilibrium, movement, sleep, autonomic functions, and even fear and aggressiveness.

During the last two decades, improved methods have been developed for inserting electrodes to a given location in the brain and fastening leads from them to the skull in such a way that the electrode remains in place for long periods of time with the skin sealed to prevent infection. It has proven possible to connect these electrodes to small radio receivers and transmitters so that electrical activity of the contacted cells can be recorded at distant receivers, and so that electrical stimulation of the contacted cells can be produced by activation of distant transmitters. In monkeys electrical stimulation of the inferior part of the lateral hypothalamus evoked marked constriction of the ipsilateral (on the same side of the body) pupil, while stimulation of another point situated six millimeters higher evoked dilation of that pupil. The magnitude of the effect was proportional to the electrical current. Simultaneous stimulation of both areas resulted in an effect that was proportional to the difference in the stimulating currents. The effect was modified by change of the illumination on the eye. Thus, it was shown that the normal sensory activity and the two electrically evoked activities could summate algebraically within the brain. In addition to pupil diameters, other functions controlled by the autonomic nervous systems could be directly controlled in a similar manner (116).

Simple motor responses could be evoked by stimulation of other parts of the brain. These responses were such movements as flexing of the left hind leg, moving the legs, raising or lowering the body, opening or closing the mouth, walking or lying still, turning around. Complex motor activity could

be evoked by stimulation of other parts of the brain. For example, stimulation of the rostral part of the nucleus caused a monkey repeatedly to interrupt its spontaneous activities, change its facial expression, turn its head to the right, stand up on two feet, circle to the right, walk on two feet, climb a pole on the back wall of the cage, descend to the floor, give a low tone vocalization, take a threatening attitude toward subordinate monkeys, then change to an attitude of peacefully approaching other monkeys, and finally resume the activity interrupted by the stimulation. This whole sequence was repeated every time the nucleus was stimulated. This response could not be evoked while the monkey was asleep or while it was strenuously engaged in other activity, such as escaping capture or eating while quite hungry. The same response could be evoked in the same way in other monkeys. It was extremely repeatable, in one case being repeated every minute day and night for two weeks for a total of more than 20,000 stimulations on a given monkey and always being evoked except during sleep.

(2) Direct control of behavior sequences. - Confirmation of this point of view is provided by a study in which electrodes were implanted in various locations in the vestibular nuclei or in the brainstem and locations found where a specific movement resulted for electric stimulation of each of the sites. The resulting movements for a single stimulation might be simple ones, such as grasping with the hand, flexing of the arm, closing of the mouth, etc. It was found that when electric stimuli were applied in sequence to a number of these sites, that a sequence of smooth movements was carried out. For example, one sequence of stimulation caused the awake monkey to grasp an object with its right hand and rapidly and smoothly bring the object to its mouth. Many other strategies of movement were performed upon changing the sequence of stimulation to the sites. In all cases, the response to the electrical stimulation overrode and preempted any simultaneous voluntary movement of the animal, and a movement sequence could be elicited whatever the position of the animal and whether it was asleep, or awake, or under anaesthesia. Destruction of the motor cortex produced paralysis to voluntary movement, yet it did not interfere with these electrically elicited movements. Cutting across the brainstem anterior to the electrode also did not alter the movements elicited by stimulation. Apparently the organization of these elementary movements exists even down to the level of the spinal cord (117).

Implanted electrodes have been used in humans to make studies in preparation for surgery on damaged portions of the brain. Stimulation of certain parts of the brain can cause the patient to become optimistic and animated or to become interested in sexual ideas, or to become passively compliant, or to feel pleasure. These examples raise a question, "Would it be feasible to control the behavior of a population by electrical stimulation of the brain?" Delgado discusses this and concludes that it is very unlikely, because of the existence of anatomical and physiological variability in individuals which would require that the control technique be developed by experiment on each individual, and because the technique requires quite specialized knowledge and refined skills (116).

11. Information Flow in the Nervous System

The external environment sends information to the sensory systems, which select and categorize it. Some of the less categorized information is immediately used (probably in the brainstem reticular formation along with information from the internal visceral environment via the hormonal and autonomic nervous system) in the assignment of priorities to subgoals. It is also used in the assignment of relevance of the various kinds of sensory information to the subgoal of highest priority. Better categorized sensory information (such as, from the specific thalamic nuclei) is used (probably in the thalamic reticular formation) in the assignment of priorities to subgoals. This in turn affects the assignment of relevance to different types of sensory information (probably at the brainstem reticular formation). That assignment, and the choice of subgoal, act (possibly at the limbic system) for the selection of appropriate action routines.

The routines call up actions (of muscles and glands) which change the external environment and the internal visceral environment. The latter changes are signalled via the hormonal system and the autonomic nervous system to parts of the brain (probably, chiefly to the reticular formation). This information on the effects of the actions acts together with such information from the external environment to act (probably via the limbic circuit and the cerebellum) to modify those actions so as to better achieve the routines ordered. That information is also supplied to a predictor computer (probably the association areas of the cerebral cortex) which then informs the modifier systems how to modify the actions further so that the action routines and subgoal are more nearly achieved.

The success or failure of the action routines in achieving the subgoal is signalled (probably by the hormonal system and autonomic nervous system) to the central nervous system (possibly the hypothalamic-pituitary portion) to increase or decrease the probability of selecting those action routines given similar input information and that subgoal. The selection probabilities are changed (probably at the recently used synapses) and affect the selection as well as the execution of action routines.

(a) Memory and learning. - Some mechanism is necessary to compare the pairings of sensory input and action routines with the subgoal which was to be attained to determine how successful those pairings were in attaining the subgoal. There should be some mechanism which causes the successful pairings (or associations) to be made easier to activate when next that subgoal is sought.

(It is a possible suggestion that modified action subroutines arrived at, possibly by the mechanism described in the previous section, are evaluated in terms of their success at achieving the high priority subgoals existing at that time (several minutes duration) by the alterations in visceral inputs and hormonal balances which result. These alterations are signalled to the brain, and act via the hippocampus to time-correlate the subroutines of the temporal and association cerebral cortex with the success or failure of the

actions. Possibly, these signals travel via the amygdala and septum to correlate the subroutines of the limbic cortex and sensorimotor cortex with the success or failure of the actions. Success and failure activate different parts of the hypothalamus and cause it to release different peptides to the pituitary and the bloodstream. These peptides may cause the pituitary to release corresponding hormones into the blood or may themselves act as hormones. These hormones act on the neuronal connections involved in the association of a subgoal and sensory inputs with action routines and subroutines. In case of "success" the corresponding hormone causes the neuronal connections to become easier to activate (it acts as a "printout signal"). In case of "failure" the hormone released acts to make the neuronal connections more difficult to activate. This is a form of learning, being the changing of the structure of the nervous system. The change in structure is the memory.)

Memory involves at least two different processes. Immediately following an experience there is a period of about three hours during which memory is not affected by an induced inhibition of protein synthesis in the brain. Memory for this period is called "short-term" memory. Both learning by mice of the correct turns in a maze and "short-term" memory, are normal during inhibition of protein synthesis in the brain. However, the ability to run the maze correctly is impaired after three hours as compared with that in mice having normal protein synthesis (118). This more lasting or permanent retention of learning is called "long-term" memory. It is not markedly impaired unless more than 90 per cent of the brain's protein synthesis is inhibited. (This suggests that protein synthesis acts at each neuronal connection that is changed during the formation of long-term memory. For the protein synthesis to act only on the neuronal connections involved in running the maze, there must have been a marker left at each active synapse. The marker might be some remnant of the chemical transmitter in the synapse or a slowly decaying (3 hours) change in the synaptic membranes. The hormone released by "success" acts at the marked synapse to turn on protein synthesis that alters the synapse so as to increase its effectiveness. A mechanism of learning similar to this suggestion has been considered in some detail by Griffith (119).)

The physiological-chemical possibility of such an action is bolstered by the finding in mice of a "nerve growth factor" which evokes the growth of nerve fibers from embryonic chick spinal ganglia. The amount of this nerve growth factor which evokes the appearance of a hundred axons from a single ganglion appears to be about ten molecules. Since each neurite seems to arise from a different neuron, each molecule of nerve growth factor must affect several cells (120). There are probably a number of different chemicals which alter synapses for long-term memory, each chemical acting at a different part of the nervous system. Support for this idea is the demonstration of a protein isolated from some mouse sarcomas, mouse salivary glands, and snake venom which selectively stimulates growth of sympathetic nerve cells (121). A number of studies have shown that the pituitary hormone ACTH, while having little effect on the acquisition of learning, markedly inhibits the ability to unlearn a no longer appropriate learned association. (This unlearning is called "extinction") (122, 123). (A possible interpretation is that ACTH acts

to interfere with the process initiated by "failure" which results in the decreased effectiveness of the immediately previously active synapses.) The natural glucocorticoid (a peptide), corticosterone (which is released by the adrenal cortex), and the synthetic glucocorticoid, dexamethasone, have been shown to facilitate extinction independently of their actions on the pituitary (123). Other synthetic glucocorticoids have been shown to inhibit extinction (124). These synthetic glucocorticoids are all peptide fragments, or slightly modified fractions of the ACTH molecule. These findings all are consistent with the model of learning suggested above.

12. The Ontogeny of Behavior

(a) Genetic determination of the structure of the nervous system. - It must be evident from the preceding discussions of nervous systems and hormonal systems that the structure of an organism is not random. There are specific types of cells having specific shapes and making specific connections with other specific cells. There are cells that produce specific chemicals and there are other cells which are acted on in particular ways by some specific chemicals. This point may seem to obvious to need mentioning, yet it is one that is often glossed over in attempts to understand the brain, let alone attempts to understand the behavior of the whole organism. Some people have postulated that the nerves grow in a random way and make their connections randomly, and that this is later modified by experience so as to lead to particular control circuits and particular nervous system behavior. Even if such a mechanism were the only one involved, it still leaves unanswered the question as to how experience can cause the formation or alteration of the connections between neurons so as to form specific types of control circuits. It is likely that to the extent that experience can cause the formation of specific neuronal connections, it does this by utilizing the mechanism which during embryonic growth causes the formation of the already-specified neuronal connections and circuits. In general, it now seems that the structure of the nervous system and the behavior of young animals is the result of developmental processes which may be modified only to a very limited extent by experience and learning. The functional assemblies of neurons in the nervous system are produced almost entirely by genetic mechanisms (125).

The lower on the evolutionary scale the animal, the less can the structure of its nervous system be affected by experience. Thus, insects are more difficult to train than frogs, and frogs more difficult than cats, and cats more difficult than humans. It has been found, for example, that the structure of the visual system in fish and amphibia can be altered only slightly by training and experience. On the other hand, the visual system of a young cat deteriorates if the cat is not given a full range of visual stimuli, such as light and dark patterns. The same apparently is true for young children also. What this seems to mean is that the neuronal connections concerned with the detection of visual patterns and the association of them with patterns of motor activity are completely and nearly permanently determined by the genetic information system in fish and amphibia. In cats, the proper connections are determined by the genetic system and are already formed at a very early age even without prior exposure to light (126), but will deteriorate somewhat with lack of use and will deteriorate greatly if

only part of normally competing neuronal pathways are activated (127, 128).

(b) Genetic determination of the visual system.

(1) Retinotectal connections in amphia. - Before considering how experience can affect the structure of the nervous system, it is worth considering how the genetic information system is able to determine the structure of the nervous system in those cases where it is known that experience cannot affect the structure. The best studied cases are those of the portions of the nervous system concerned with vision in lower vertebrates. The eye connects to the brain via the fibers of the optic nerve. These fibers are axons of the ganglion cells of the retina. Each of these retinal ganglion cells can be activated by application of a light pattern to some particular region of the retina; each ganglion cell connects to its own region of the retina, called its receptive field. These receptive fields usually overlap the neighboring ganglion cells and may vary in visual angle from 1° to 40° depending on the type of ganglion cell. In fish and in amphibia, the majority of these ganglion cell axons terminate on a part of the brain called the optic tectum. It is found that these optic nerve terminations on the tectum occur in a layer parallel to the surface of the tectum and with a topographical arrangement corresponding to the topographical arrangement on the retina of the ganglion cell bodies from which they originate. This topographical mapping on the brain is quite consistent from one individual to another in the same species. In the case of the frog, it has been found that there are five main types of ganglion cells that terminate on the optic tectum. These five types differ in the particular light pattern which activates them. Two of the types terminate in the same layer parallel to the tectal surface and each of the three remaining types in its own of three layers parallel to this layer. What is especially interesting is the finding that the topographical maps thus formed for the five types of ganglion cells are all congruent to each other in the tectum.

Within the tectum, these optic nerve terminations contact a number of different types of tectal neurons, including some which receive inputs from other sense organs. Their outputs are transformed and sent to another layer of tectal cells which connect to various motor nuclei, and control patterns of muscle actions. A frog is capable without training of using visual information concerning a fly passing it to control the direction and distance of its jump and coordinate this with the opening of its mouth and the flicking out of its tongue so as to wrap its tongue around the fly. This requires that all the neuronal connections from the retina to the tectum, through the tectum to the motor nuclei, and to the various muscles be so constructed as to ensure this coordination between the visual pattern and the resulting muscle activity.

(2) Determination of optic fiber terminations at the tectum. - As a starting point, let us consider how the topological information about the location of the retinal ganglion cells in the retina is preserved in the location of their axonal terminations in the optic tectum. This cannot be explained as a result of the axons always keeping the same relation to each other that their cell bodies had and thus ending up in the same relations at their terminations. It has been found by electron microscopy that the axons

within the optic nerve twist and wind around each other very much like the threads in a braid. Furthermore, this can be shown by another type of experiment. If the optic nerves are cut in fish or amphibia, the parts of the axons separated from the cell bodies (those portions that terminate at the tectum) will degenerate. The cut ends of the axons which are still connected to the cell bodies will begin to grow and will first spread out in every which direction and form a sort of tangled knot of axons. These then begin growing in the general direction of the tectum, divide into the two main branches that go to the top and bottom on the tectum and then spread out along the tectum, making what at first seem to be randomly located contacts. In a few weeks though, those contacts are sorted out and in the majority of cases the normal mapping of the optic fiber terminations on the tectum is reestablished. This indicates that whatever factors operated during the original development of the visual system are still operating during the regeneration of the optic nerve. This regeneration of normal visual connections in amphibia was established first by the observation that the visuomotor behavior was fully restored (128a), and later confirmed by mapping the retinotectal projection electrically at the terminations of the optic nerve fibers in the tectum (129, 130). Whatever these controlling factors are, they must have some way of recognizing when an optic fiber has reached its proper location in the tectum. This implies that the factors must also be able to recognize each optic fiber. In other words there must be some specificity of each optic nerve fiber which is matched in some way by a specificity of the tectal cells in the location of its proper termination.

(3) Tectal connections of modified retinas. - The first evidence to consider is presented in a series of papers (131, 132, 133) dealing with the retinotectal projection in Xenopus, an amphibian. The normal projection is from an eye to the tectum on the opposite side of the brain. The part of the retina farthest from the nose (temporal) projects to the part of the tectum closest to the front of the head (rostral); the part of the retina closest to the nose (nasal) projects to the part of the tectum closest to the rear of the head (caudal). Thus, the temporal-nasal axis in the retina is projected onto a rostral-caudal axis on the tectum. In a similar way, the lower part of the retina (inferior) projects to the part of the tectum furthest from the midline of the head (lateral); and the top of the retina (superior) projects to the part of the tectum closest to the midline (medial). Thus, there is also another axis of projection at right angles to the preceding; the inferior-superior axis of the retina projects to the lateral-medial axis of the tectum. To test what controls the specificity of these projections, in embryonic animals after the eyes began to form they were split in half and half the eye was removed and replaced by a half eye from another animal; this grafted half then growing into place. Some animals were prepared with eyes in which the temporal halves of the retinas had been replaced with nasal halves. In these double-nasal retinas, as one moved temporally outward along the nasal-temporal axis, the retinas corresponded to nasal-central-nasal instead of the normal nasal-central-temporal. In a similar way, double-temporal retinas were prepared in other animals. All these operations were done before the optic fibers from the retinas had yet made connections with the optic tectum. These animals were then raised until they were adults, and electrophysiological techniques were used to map

the retinotectal projections. The findings were that in animals with double-temporal retinas, the visual projection from each half-retina covered the whole tectum. This was so also in animals with double-nasal retinas. In each case, the pattern of projection was retinotopically organized in a way that was appropriate to the original half-retina, and was a mirrored image of this pattern for the grafted half-retina. Each point on the tectum received nerve impulses from two positions on the retina, and these were symmetrically disposed about the vertical meridian. The inferior-superior axis of the retina projected in a normal way onto the lateral-medial axis of the tectum. These results show the independence of the nasal-temporal axis and the inferior-superior axis.

The results could be interpreted as indicating a gradient of specificity along the temporal-nasal retina which was matched by a gradient of specificity along the rostral-caudal tectum. Thus, each optic fiber that originated from a position a certain fraction of the way along the temporal-nasal gradient of a half retina projected onto a tectal point corresponding to the same fraction of the rostral-caudal gradient of the whole tectum. In a similar way, it could be assumed that there was an independent gradient at right angles which determined the connections of the inferior-superior axis of the retina to the lateral-medial axis of the tectum. Two possible mechanisms have been suggested which would give the observed connections of the compound eye (double half-retinas) with the optic tectum. First, that the optic fibers from each (similar) half of the compound eye connect only with their appropriate half of the tectum; the remaining half of the tectum failing to develop. Second, the optic fibers from each half retina spread out to connect with the whole tectum. It was found by measurement that the optic fibers from the medial-lateral axis had approximately the normal spread, while along the rostral-caudal axis they had about twice the normal spread. This result is compatible with the first mechanism of the spreading of the optic nerve termination. It is also compatible with the second mechanism if it is assumed that the remaining half of the layer of tectal cells with which the optic fibers connect spreads out so as to fill the whole surface.

(4) Formation of gradients of specificity in the retina. - We will not concern ourselves for the moment with the mechanism by which the specific termination of the optic fiber of the tectum is produced, but only with the mechanisms which could lead to the formation of a pair of orthogonal independent gradients of specificity on the retina which in some way matches a pair of orthogonal gradients on the tectum. To do this, we must look at the embryological development of the retina and the tectum. The organism starts as a single fertilized egg cell. This cell does not have a uniform distribution of chemicals within it, and corresponding to this asymmetry there is an asymmetrical division of the cell in two, a further asymmetrical division of the daughter cells into four, and so on. This results in an asymmetrical hollow sphere with a wall consisting of a single layer of cells. An indentation forms at one pole of the sphere. This indentation becomes very deep and at that point is called the archenteron. Cells move across the wall of the sphere and enter the archenteron and locate themselves at its roof. These cells have been found to be the cells having the greatest power to induce

organization and differentiation in neighboring cells. These cells in the roof of the archenteron induce the formation of the neural plate in the overlying outer wall (the ectoderm). The cells making up the archenteron roof have by now differentiated into a type of cell called the mesoderm from which will later arise muscles. The strip of archenteron roof forms an axis on either side of which is a bilateral symmetry. This could be called the lateral-medial gradient. Cells move along the archenteron from the opening (posterior) to the back (anterior) of the invagination. These cells seem to carry along at least two different substances which induce changes in the overlying archenteron roof and neural plate. The first substance to enter on reaching the deepest portion of the archenteron (anterior) induces the cells of the neural plate to start differentiating into what later will become the brain, eyes, and nasal organs. A second substance enters and acting at the posterior portion of the archenteron, induces the formation of the spinal cord in the neural plate and causes the mesoderm of the archenteron roof to start forming the structures of the posterior trunk and tail.

The neural plate forms lateral ridges which roll up and close to make a tube, the neural tube. At the anterior end of the tube, from the left and right sides, protrusions come out which are the primary eye vesicles. The protrusions constrict at the ends nearest the neural tube to become the eyestalks with ball-like eye-vesicles at their lateral end. The outer wall of the eye-vesicle finally establishes contact with the skin ectoderm at the side of what has by now developed into the head. This induces the skin ectoderm to thicken, forming a group of cells called the lens placode. The lens placode detaches from the ectoderm and moves toward the eye-vesicle and as it does the eye-vesicle invaginates away from the placode, forming a double-walled eye cup. Some of the mesoderm cells surrounding it penetrate into the eye cup to help form what later becomes a vitreous body and the several tough membranes around the eye. As the lens placode comes in contact with the inner of the two walls of the eye cup, the inner wall thickens and develops into the retina, while the outer wall forms into the pigment epithelium. Later the iris arises from the edges of the eye cup; the axonal fibers grow from the retinal ganglion cells down through the eyestalk to form the optic nerve and eventually connect with the optic tectum. The skin overlying the eye forms into the transparent cornea.

The potency for formation of an eye-vesicle in the neural tube is induced in ectoderm by the anterior part of the archenteron roof before the neural plate forms. After a certain state of development, a narrow median strip of the archenteron roof induces in the overlying ectoderm a local suppression of the eye-forming potency. Thus, the eye-vesicles form at two sides rather than at the center of the neural tube. Occasionally the median strip fails to form and then a single central eye develops (a cyclops). The continuing differentiation of neural ectoderm into retina and pigment epithelium is also established at this time as a potential ability, which is brought into action only when in contact with the lens placode. The archenteron roof also is responsible for a weak lens-placode-forming potency of the skin ectoderm. This potency is increased when the eye-vesicle contacts the ectoderm, and results in the formation of the lens placode.

Note that the archenteron roof, which has a medial-lateral gradient, acquires an anterior-posterior gradient during invagination. The archenteron roof induces the lens-forming potency in the anterior skin ectoderm and possibly at the same time establishes anterior-posterior and medial-lateral gradients in the skin. The archenteron roof also induces the eye-vesicle-forming potency, the specialization of the anterior neural ectoderm and the topography of the different portions of the brain, and possibly induces in all of these corresponding anterior-posterior and medial-lateral gradients (134, 135).

It has been shown by Stone (136) that the eye-vesicle can be rotated within the head without affecting the establishment of normal connections between the subsequently developed retina and optic tectum right up to the time that the optic vesicle is contacted by the lens placode. After that time, rotation of the optic vesicle will cause the development of an eye that is rotated in the head with respect to its connections to the brain. From this it seems reasonable to conclude that it is the polar gradients of the placode which induce in the eye-vesicle its final gradients of specificity.

(5) Determination of specific connections within the retina. - When the lens placode comes in contact with the eye-vesicle, it causes the cells of the inner wall to start dividing. Some of their daughter cells differentiate into the ganglion cells. Later, the sister cells differentiate into the bipolar cells and horizontal cells and amacrine cells and finally into the visual cells (rods and cones) of the retina. It is at the time that the ganglion cells have differentiated that the polarization of the retina, as measured by the specificity of the optic fiber terminations, is established. Experiments have shown that the polarization in the anterior-posterior axis of the eye cup becomes fixed first. This is at so-called stage 34 in the Xenopus; and at stages 35 and 36 of development, the dorsal-ventral axis becomes fixed in its polarity. It is not until stage 38 that the optic fibers connect to the tectum. In the salamander, the anterior-posterior axis is established at stage 30, and five to ten hours later, at stage 31, the dorsal-ventral axis of the retina is established (133). Stone also tried experiments in which he rotated, at various stages of development the piece of neural tube which would become the optic tectum while leaving the eye in its normal position. He found that polarization of the optic tectum so as to cause rotation of the connections between eye and tectum occurred between stages 31 and 36. The connections from the retina to the tectum are not formed until later, at about stage 38. It has been mentioned in the previous discussion of neuronal circuits in the frog's retina how there are different types of ganglion cells and each makes specific connections with certain of the amacrine cells and bipolar cells and these, in turn, make specific connections with only certain of each other, and the visual cells in turn make connections only with specific of the bipolar cells and horizontal cells. The specificity of these connections may be determined by a mechanism in which the specifically connecting set of ganglion cells and other retinal cells all derive from the same mother cell of the inner wall of the eye-vesicle and thus all retain enough similarity of chemical specificity to insure the specificity of their interconnections.

(6) Mechanism of termination selection along a gradient. - As has been mentioned above, the optic fiber connections along the anterior-posterior axis of the optic tectum are specified before the connections are specified along the dorsal-ventral axis. The anterior-posterior axis later becomes the rostral-caudal axis and the dorsal-ventral axis becomes the medial-lateral axis. One mechanism that has been proposed to explain the determination of these connections along each axis separately is that there is a gradient of some particular chemical along the temporal-nasal axis of the retinal ganglion cells which corresponds to a rostral-caudal concentration gradient of the matching chemical at the surface of the optic tectum (possibly in the dendrites of the tectal cells to which the optic fibers connect). It is also assumed that there is a second pair of gradients of a different pair of matching chemicals along the inferior-superior axis of the retina and the lateral-medial axis of the optic tectum. This latter pair of chemical concentration gradients is presumed to be formed a few hours after the first. It is observed that the optic nerve fibers first make connections at random on the tectal surface, but that gradually the fibers redistribute themselves until they are properly lined up along, first, the anterior-posterior axis and then afterwards along the dorsal-ventral axis. It has been noted that the connections unfold from the rostral (anterior) to the caudal (posterior) poles of the tectum. This parallels the direction of development and differentiation of the cells of the optic tectum. It is assumed that the terminations of the optic nerve fibers carry the same chemical concentrations as the parent cell bodies. There is considerable evidence that many types of chemicals are passed from the nerve cell body along the axons to the terminations. Thus, it has been shown that protein is moved along the axons of neurons (137). It is assumed that the final connection of an optic fiber along the concentration gradient of the tectum is determined by the concentration of the chemical in that fiber with respect to the concentration in the neighboring fibers. (A possible mechanism is that the rate of formation of connections between the optic fiber and the tectal cells may depend on the sum of the concentrations of the specific chemicals in the optic fiber and in the tectal cell with which it is making contact. Thus, the optic fibers having the higher concentration of the chemical will tend to displace the optic fibers having lower concentrations and will simultaneously tend to move toward the higher concentration regions of the tectum. This will automatically line up the low to high concentration fibers along the low to high concentration axis of the tectum. The existence of such a mechanism is consistent with the findings; first, that the nerve fibers which normally connect to the cells of a particular muscle will displace foreign nerve fibers that have already connected to those muscle cells (138, 139). Second, when the optic fibers are first lined up on the tectum, the axons have an extensive overlap of their fine terminal connections on the tectal cells. Later this overlap is reduced and the terminations become more selective.)

(7) Cell-to-cell specificity. - Once the gradients and corresponding connections have been established, the tectal cells and the optic nerve axons take on specific values which match each to the other. Thereafter, connections can be made only between the corresponding axon and tectal cell. This is shown by removing a piece of retina or a piece of tectum at these later stages, and then cutting the whole optic nerve. When the optic nerve

regenerates, connections are not made to the parts of the tectum corresponding to the missing pieces of the retina, and connections from the retina are not made to the tectum if the corresponding piece of tectum is missing. The same thing has been found to occur for the retinotectal connections of the chick (140, 144).

A molecular mechanism for the initial polarization of tissues, which is based on mechanisms found to work on bacterial cells, is discussed in the literature (142). A more detailed discussion and elaboration of this possible mechanism is given in chapter three of reference (119). Other mechanisms have been discussed in section 2 of this report. Further evidence that such a molecular mechanism of genetic control of the specificity of cells is operating in the retina is given in some work (143) in which it was found for Xenopus that DNA synthesis ceased in the retinal ganglion cell neuroblasts (cells which later become neurons) at stages 28-29, and that the specification of the tectal connections of the ganglion cells is produced at stages 29-31. There is less than a ten-hour period between stages 28 and 30, which is less than one generation cycle for the retinal cells. The retinal ganglion cell neuroblasts can change their specificity of tectal connection to match the position of the eye-vesicle in the head up to the time that DNA replication ceases and cannot change their specificity after that time. Therefore, it is consistent with the idea that neuronal specification involves the synthesis of specific macromolecules, synthesis of which is controlled by DNA.

(8) Tectal output connections. - Some idea as to specification of neurons within the optic tectum is given by experiments performed on amphibia (144). Animals which were given double-temporal or double-nasal eyes during embryonic development later behaved in their visual pursuit reactions as if the entire visual field were nasal or temporal, respectively. After destruction of the rostral half of the tectum to which the temporal retina (nasal field) normally projects, the animals with double-temporal retinas (double-nasal visual field) behaved as if they were totally blind. Similarly, animals with double-nasal retinas (double-temporal visual fields) became blind when the caudal half of the tectum was destroyed. Since it is known that both types of double-half eyes had optic fiber projections to the entire surface of the optic tectum, the question arises as to why the animal should become totally blind when just one half of the tectum was destroyed.

(A possible suggestion is that the nerve cells of the tectum which connected to the motor nerves that activated turning in the temporal direction are located in the caudal portion of the tectum, and that the corresponding tectal cells for turning in the nasal direction are located in the rostral portion of the optic tectum. These cells make dendritic connections either directly, or more likely via other tectal neurons, with the terminations of optic fibers in the tectal surface. A double-temporal eye, for example, would have its optic fiber termination spread across the entire rostral-caudal axis of the tectal surface. The tectal cells that normally connected with the optic fibers from the temporal half of the retina would spread their dendritic terminations across the entire tectal surface to connect with those optic fibers. Or possibly, if there was an intervening layer of connecting tectal cells, all of these outer tectal cells would make connections with the fibers

terminating at the normal positions of these cells in the entire tectal surface, but only those cells which normally would connect to the temporal half of the retina would in turn get dendritic connections from the output tectal cells. Thus, the output tectal cells connecting to the motor neuron would be innervated only in the rostral half of the tectum for a tectum connecting to a double-temporal eye and only in the caudal half of the tectum for a tectum connecting to a double-nasal eye. Destruction of the rostral half of the tectum for the double-temporal eyes would cause total functional blindness because the incoming visual signal to the tectum could not cause any corresponding activation of motor neurons. Similarly, destruction of the caudal half of the tectum would make an animal with double nasal eyes functionally blind.)

(c) Other examples of specific neuronal connections and their control.

(1) Control of sensory nerve types. - It has been found that sensory nerves from the skin of the frog's head terminate medially in the medulla (a region of the brainstem), while sensory nerve fibers from the cornea of the frog's eye terminate laterally in the medulla. The fibers from the cornea are much smaller in diameter and slower in conduction rate than those from the skin of the head. The laterally terminating sensory fibers appear to make a monosynaptic connection in the medulla with the dendrites of those spinal motor neurons which activate the muscles that produce a blinking of the eyelids. Whenever the cornea was stimulated, a blink reflex occurred. If an extra eye-vesicle was grafted onto a frog embryo in the head region, and the frog grown to an adult with an extra eye, the sensory nerve fibers connecting to the piece of skin which had changed into cornea were of small diameter instead of the large diameter which they would normally have been. These small diameter fibers from the newly formed cornea were found to terminate laterally in the medulla. It is known from other work that if sensory nerve fibers have their dendritic terminations in tissue which is very poorly supplied with blood, the nerve fibers will become small in diameter. The cornea is a tissue which is lacking a capillary blood supply, and therefore would be expected to cause the cutaneous sensory nerves to become small in diameter. Apparently, the chemical changes which are part of this change in diameter cause the central axonal end of the sensory fiber to terminate laterally rather than medially in the medulla and thus to contact the proper motoneuron dendrites to set up the blink reflex (145, 146).

It has been found in experiments with adult frogs that the sensory nerves going to the skin and tongue could be cut and the central stump of the cutaneous (skin) nerve allowed to grow to the tongue in place of the normal nerve. It was found (147) that when regeneration was completed, the cutaneous nerves cross-innervating the tongue often yielded electrical responses which were "gustatory" by all criteria, responding in a characteristic way to stimulation of taste receptors in the tongue. Thus, cutaneous nerves can serve the same role as the gustatory nerves. Therefore, the functional characteristics of these sensory nerve cells are not predetermined, but must depend on the environment of the dendritic endings of the neurons. This, taken together with the evidence described in the preceding paragraph, indicates that it is the skin type (with its corresponding receptors) which determines the character-

istics of a sensory neuron that synapses with the receptors in each piece of skin. The type of neuron, in turn, determines where its axon synapses centrally in the spinal cord or brain.

A third example of the specification of nerves by the skin to which they connect is the results of a study of frogs which when tadpoles had had skin grafts rotated 180° (148). When the normal skin was tickled, a leg moved to scratch at the point of stimulation. For normal skin and for grafted back skin that was rotated head-to-tail the reflex movements were accurately directed. For stimulation of grafted belly skin on the back, the leg wiped the belly, not the back. Similarly, for back skin on the belly, the back was wiped. The nerves to the skin receptors were mostly the local nerves and each entered the skin within its own receptive field, so it cannot be that nerves not normally found in that part of the body had grown to the grafted skin. Furthermore, when the reflexes first appeared, they were often normal, even in the grafted skin. Only later did the reflexes become misdirected. The receptive fields of the cutaneous nerves remained unchanged, so it must have been the central connections that changed (149). The conclusion is that the cutaneous nerves were modified by the skin with which they connected so as to specify their spinal cord connections with motoneurons and with other central neurons to be appropriate for the skin type rather than for the position at which the nerve ended in the body's skin.

(2) Control of muscle cell type by motoneurons. - The higher vertebrates have two types of skeletal muscle fiber or cell. The fast fibers give a twitch contraction upon a single neuronal stimulation. The slow fibers give a much slower mechanical response and require repetitive stimulation of their motor nerve fibers to give a good mechanical response. The slow muscle fibers are innervated by nerve fibers of smaller diameters than those innervating the fast muscle fibers. The slow muscle fibers have multiple nerve endings on each muscle cell, while the fast muscle fibers have a single motoneuron synapse on each cell. Some skeletal muscles have only slow muscle fibers, some have only fast muscle fibers, and some have a mixture of the two types of muscle fibers. It was found in experiments on kittens and rabbits (150), rats (151), and chicks (152) that when the motor neurons that normally innervated slow muscles were cut and regrown to fast muscles, the types of nerve endings on the muscle fibers became those of the new type of motoneuron (slow) after regeneration of the nerve. Furthermore, the muscle fiber showing only the electrical and mechanical activity characteristic of slow muscle fibers, was shown to have changed in structure to that characteristic of slow fibers, and was shown to have changed its content of enzymes to those characteristic of the slow fibers. In a similar way, connection of the motoneurons that normally innervated fast muscles to slow muscles caused the slow muscle fibers to change in all these ways to fast muscle fibers. It must be concluded that the metabolic system, enzyme concentration, resulting structure and consequent electrical and mechanical action of the skeletal muscle fibers are determined by the nerve supply. It is not yet known whether the determination is by the type of electric activity from the nerve terminations or by substances transmitted from the nerve terminations to the muscle fiber.

A book which discusses in detail the development of neuronal connections to the limbs of amphibians has just become available (153).

(d) Overall view of the problem of neuronal specificity. - We can take the view that the patterns of environmental stimulation reaching the surface of the animal must be matched with the pattern of response (muscle actions) taken by the animal. This means that information as to what portions of the skin were stimulated must somehow be retained so that the appropriate muscles can, in turn, be stimulated. The matching of the input and the output patterns is done at the connections in the central nervous system so that information as to the input origins and the output terminations must somehow be preserved in the central nervous system connections. As was pointed out above, the polarizations of the outer skin and of the central nervous system were both produced by the archenteron roof during the early embryonic development of the animal. Thus, matching polarities are set up which can insure proper connections from the sensory receptors at the outer skin via nerves (which are changed in character by the skin from which they originate) to terminations of those nerves at the matching portions of the central nervous system. Not much seems to be known about the formation of connections between the central nervous system and the muscles. However, one can speculate that since the archenteron roof tissue is itself the source of the cells that become the muscles, the muscles themselves retain a polarization that matches them to the polarization of the central nervous system. Thus, the nerve fibers going from the central nervous system to the muscles would match up with the corresponding muscles. The particular type of the muscle would then be determined by the nerve fibers that terminated on it. In this way, a matching of connections from outer skin to central nervous system to output motoneurons to muscles is retained, and the overall input and output patterns are matched to each other. Naturally, further experiments are needed to test this hypothesis.

Bibliography

1. Kuyper, Ch. M. A., "The Organization of Cellular Activity", Elsevier, Amsterdam and N. Y., 1962.

2. Bonner, D. M. (Editor), "Control Mechanisms in Cellular Processes", Ronald Press Co., N. Y., 1960.

3. Umbarger, H. E., End-product Inhibition of the Initial Enzyme in a Biosynthetic Sequence as a Mechanism of Feedback Control, pp. 67-86 in Ref. 2, 1960.

4. Van Eys, J., Regulatory Mechanisms in Energy Metabolism, pp. 141-166 in Ref. 2, 1960.

5. Vogel, H. J., Control by Repression, pp. 23-66 in Ref. 2, 1960.

6. Bretscher, M. S., How Repressor Molecules Function, Nature, 217, 509, 1968.

7. Mitchell, P., Proton-Translocation Phosphorylation in Mitochondria, Chloroplasts and Bacteria; Natural Fuel Cells and Solar Cells, Fed. Proc., 26, 1370, 1967.

8. Chance, B., Bonner, Jr., W. D., Storey, B. T., Electron Transport in Respiration, Ann. Rev. Plant Physiol., 19, 295, 1968.

9. Chance, B., Mela, L., Harris, E. J., Interaction of Ion Movements and Local Anesthetics in Mitochondrial Membranes, Fed. Proc., 27, 902, 1968.

10. Suskind, S. R., Yanofsky, C., Genetic Control of Enzyme Structure, pp. 3-22 in Ref. 2, 1960.

11. Jacob, F., Monod, J., Genetic Regulatory Mechanisms in the Synthesis of Proteins, J. Mol. Biol., 3, 318, 1961.

12. Spangler, R. A., Snell, F. M., Transfer Function Analysis of Chemical Kinetic Systems, J. Theor. Biol., 16, 366, 1967.

13. Spangler, R. A., Snell, F. M., Transfer Function Analysis of an Oscillatory Model Chemical System, J. Theor. Biol., 16, 381, 405., 1967.

14. Hastings, J. W., Temporal Regulation in Cellular Processes, pp. 227-243 in Ref. 2, 1960.

15. Pittendrigh, C. S., Bruce, V. S., Daily Rhythms as Coupled Oscillator Systems and Their Relation to Thermoperiodism and Photoperiodism. In "Photoperiodism and Related Phenomena in Plants and Animals", Withrow, R. B. (Editor), Am. Assoc. Advancement of Sci. Publication No. 55, Washington, D. C., 1959.

16. Goodwin, B. C., "Temporal Organization in Cells, A Dynamic Theory of Cellular Control Processes", Academic Press, London and N. Y., 1963.

17. Herrmann, H., Direct Metabolic Interactions Between Animal Cells, Science, 132, 529, 1963.

18. Grobstein, C., Cytodifferentiation and Its Controls, Science, 143, 643, 1964.

19. Tomkins, G. M., et al, Control of Gene Expression in Higher Organisms, Science, 166, 1474, 1969.

20. Moscona, M. H., Moscona, A. A., Inhibition of Adhesiveness and Aggregation of Dissociated Cells by Inhibitors of Protein and RNA Synthesis, Science, 142, 1070, 1963.

21. Jacobson, A. G., Inductive Processes in Embryonic Development, Science, 152, 25, 1966.

22. von Bertalanffy, L., "Modern Theories of Development", Oxford Univ. Press, 1933. Also Harper Torchbooks, 1962.

23. Orr, C. W., Roseman, S., Intercellular Adhesion. I. A Quantitative Assay for Measuring the Rate of Adhesion, J. Membrane Biol., 1, 109, 1969.

24. Orr, C. W., Roseman, S., Intercellular Adhesion. II. The Purification and Properties of a Horse Serum Protein that Promotes Neural Retina Cell Aggregation, J. Membrane Biol., 1, 125, 1969.

25. Overton, J., A Fibrillar Intercellular Material Between Reaggregating Embryonic Chick Cells, Expl. Cell Res., 40, 136, 1969.

26. Flickinger, R. A., Cell Differentiation: Some Aspects of the Problem, Science, 141, 608, 1963.

27. Britten, R. J., Davidson, E. H., Gene Regulation for Higher Cells: A Theory, Science, 165, 349, 1969.

28. Kanno, Y., Loewenstein, W. R., Low-Resistance Coupling Between Gland Cells. Some Observations on Intercellular Contact Membranes and Intercellular Space, Nature, 201, 194, 1964.

29. Loewenstein, W. R., Socolar, S. J., Higashino, S., Kanno, Y., Davidson, N., Intercellular Communication: Renal, Urinary, Bladder, Sensory, and Salivary Gland Cells, Science, 149, 295, 1965.

30. Kanno, Y., Loewenstein, W. R., Cell-to-Cell Passage of Large Molecules, Nature, 212, 629, 1966.

31. Nakas, N., Higashino, S., Loewenstein, W. R., Uncoupling of an Epithelial Cell Membrane Junction by Calcium-Ion Removal, Science, 151, 89, 1966.

32. Wiener, J., Spiro, D., Loewenstein, W. R., Studies on an Epithelial (Gland) Cell Junction, II. Surface Structure, J. Cell Biol., 22, 587, 1964.

33. Bullivant, S., Loewenstein, W. R., Structure of Coupled and Uncoupled Cell Junctions, J. Cell Biol., 37, 621, 1968.

34. Loewenstein, W. R., Nakas, M., Socolar, S J., Junctional Membrane Uncoupling: Permeability Transformations at a Cell Membrane Junction, J. Gen. Physiol., 50, 1865, 1967.

35. Loewenstein, W. R., Cell Surface Membranes in Close Contact. Role of Calcium and Magnesium Ions, J. Colloid Interface Sci., 25, 34, 1967.

36. Loewenstein, W. R., On the Genesis of Cellular Communication, Develop. Biol., 15, 503, 1967.

37. Loewenstein, W. R., Permeability Conversion at a Membrane Junction, pp. 234-238 in Bolis, L., Pethica, B. A. (Editors), "Membrane Models and the Formation of Biological Membranes", North-Holland Publishing Co., Amsterdam, 1968.

38. Politoff, A. L., Socolar, S. J., Loewenstein, W. R., Permeability of a Cell Membrane. Dependence on Energy Metabolism, J. Gen. Physiol., 53, 498, 1969.

39. Ito, S., Loewenstein, W. R., Ionic Communication Between Early Embryonic Cells, Develop. Biol., 19, 228, 1969.

40. Loewenstein, W. R., III. Emergence of Order in Tussues and Organs. Communication through Cell Junctions. Implications in Growth Control and Differentiation, Develop. Biol. Suppl., 2, 151, 1968.

41. Jamakosmanović, A., Loewenstein, W. R., Intercellular Communication and Tissue Growth. III. Thyroid Cancer, J. Cell. Biol. 38, 556, 1968.

42. Loewenstein, W. R., Kanno, Y., Intercellular Communication and Tissue Growth. I. Cancerous Growth, J. Cell. Biol., 33, 225, 1967.

43. Rose, B., Loewenstein, W. R., Depression of Junctional Membrane Permeability by Substitution of Li^+ for extracellular Na^+, Biochim. Biophys. Acta 173, 146, 1968.

44. Wolpert, L., Positional Information and the Spatial Pattern of Cellular Differentiation, J. Theoret. Biol., 25, 1, 1969.

45. Saunders, J. W., Gasseling, M. T., Trans-filter Propagation of Apical Ectderm Maintenance Factor in the Chick Embryo Wing Bud, Develop. Biol., 7, 64, 1963.

46. Goodwin, B. C., A Phase-Shift Model for the Spatial and Temporal Organization of Developing Systems, J. Theoret. Biol., 25, 49, 1969.

47. Winfree, A. T., Biological Rhythms and the Behavior of Populations of Coupled Oscillators, J. Theoret. Biol., 16, 15, 1967.

48. Robertson, J. D., The Membrane of the Living Cell, Sci. Amer., 206, 64, 1962.

49. Satir, P., Ciliated Cells, Sci. Amer., 204, 108, 1961.

50. Loewenstein, W. R., Facets of a Transducer Process, Cold Spring Harbor Symp. Quant. Biol., 30, 29, 1965.

51. Loewenstein, W. R., Excitation and Inactivation in a Receptor Membrane, Ann. N. Y. Acad. Sci., 94, 510, 1961.

52. Lipetz, L. E., The Transfer Functions of Sensory Intensity in the Nervous System, Vision Res., 9, 1205, 1969.

53. Smith, T. G., Stell, W. K., Brown, J. E., Conductance Changes Associated with Receptor Potentials in Limulus Photoreceptors, Science, 196, 456, 1968.

54. Smith, T. G., Stell, W. K., Brown, J. E., Freeman, J. A., Murray, G. C., A Role for the Sodium Pump in Photoreception in Limulus, Science, 162, 456, 1968.

55. Wersäll, J., Flock, A., Physiological Aspects of the Structure of Vestibular End Organs, Acta Oto-laryng., Suppl. 92, 85, 1965.

56. Szabo, T., Sense Organs of the Lateral Line System in Some Electric Fish of the Gymnotidae, Mormyridae and Gymnarchidae, J. Morph., 117, 229, 1965.

57. Davis, H., A Model for Transducer Action in the Cochlea, Cold Spring Harbor Symp. Quant. Biol., 30, 181, 1965.

58. Bodian, D., The Generalized Vertebrate Neuron, Science, 137, 323, 1962.

59. Bishop, G. H., Natural History of the Nerve Impulse, Physiol. Rev., 36, 376, 1956.

60. Edwards, C., Otooson, D., The Site of Impulse Initiation in a Nerve Cell of a Crustacean Stretch Receptor, J. Physiol., 143, 138, 1958.

61. Katz, B., Depolarization of Sensory Terminals and the Initiation of Nerve Impulses in the Muscle Spindle, J. Physiol., 111, 261, 1950.

62. MacNichol, Jr., E. F., Visual Receptors as Biological Transducers, In "Molecular Structure and Functional Activity of Nerve Cells", Publication No. 1 of AIBS, 34, 1956.

63. Grundfest, H., Synaptic and Ephaptic Transmission, Chap. V of "Handbook of Physiology, Sec. 1, Neurophysiology", Vol. I, Amer. Physiol. Soc., 1959.

64. Dudel, J., Kuffler, S. W., Mechanism of Facilitation at the Crayfish Neuromuscular Junction, J. Physiol., 155, 530, 1961.

65. Kandel, E. R., Frazier, W. T., Coggeshall, R. E., Opposite Synaptic Actions Mediated by Different Branches of an Identifiable Interneuron in Aplysia, Science, 155, 346, 1967.

66. Gerschenfeld, H. M., Ascher, P., Tauc, L., Two Different Excitatory Transmitters Acting on a Single Molluscan Neuron, Nature, 213, 358, 1967.

67. Wachtel, H., Kandel, E. R., A Direct Synaptic Connection Mediating both Excitation and Inhibition, Science, 158, 1206, 1967.

68. Katz, B., Miledi, R., Input-Output Relation of a Single Synapse, Nature, 212, 1242, 1966.

69. Ochs, S., "Elements of Neurophysiology", John Wiley & Sons, N. Y., 1965.

70. Maffei, L., Rizzolatti, G., Transfer Properties of the Lateral Geniculate Body, J. Neurophysiol., 30, 333, 1967.

71. Rall, W., Theoretical Significance of Dendritic Trees for Neuronal Input-Output Relations, In Reiss, R. F. (Editor), "Neural Theory and Modeling", Stanford Univ. Press, pp. 73-97, 1964.

72. Purple, R. L., Dodge, F. A., Interaction of Excitation and Inhibition in the Eye of Limulus, Cold Spring Harbor Symp. Quant. Biol., 30, 529, 1965.

73. Hartline, H. K., Ratliff, F., Inhibitory Interaction of Receptor Units in the Eye of Limulus, J. Gen. Physiol., 40, 357, 1957.

74. Eccles, J. C., "The Physiology of Synapses", Academic Press, N. Y., 1964.

75. Howland, B., Lettvin, J. Y., McCulloch, W. S., Pitts, W. H., Walls, P. D., Reflex Inhibition by Dorsal Root Interaction, J. Neurophysiol., 18, 1, 1955.

76. Dudel, J., Presynaptic Inhibition of the Excitatory Nerve Terminal in the Neuromuscular Junction of the Crayfish, Pflügers Arch. ges. Physiol., 277, 537, 1963.

77. Armett, C. J., Gray, J. A. B., Hunsperger, R. W., Lal, S., The Transmission of Information in Primary Receptor Neurones and Second-Order Neurones of a Phasic System, J. Physiol., 164, 395, 1962.

78. Van Bergeijk, W. A., Studies with Artificial Neurons: II. Analog of the External Spiral Innervation of the Cochlea, Kybernetik, 1, 102, 1961.

79. Crane, H. D., Possibilities of Signal Processing in Axon Systems, In Reiss, R. F. (Editor), "Neural Theory and Modeling", pp. 138-153, Stanford Univ. Press, 1964.

80. Nomoto, M., Suga, N., Katsuki, Y., Discharge Pattern and Inhibition of Primary Auditory Nerve Fibers in the Monkey, J. Neurophysiol., 27, 768, 1964.

81. Dowling, J. E., Boycott, B. B., Organization of the Primate Retina: Electron Microscopy, Proc. Roy. Soc., B166, 80, 1966.

82. Ratliff, F., Inhibitory Interaction and the Detection and Enhancement of Contours, In Rosenblith, W. A. (Editor), "Sensory Communication", pp. 183-203, MIT Press and John Wiley & Sons, 1961.

83. Frazier, W., Waziri, R., Kandel, E. R., Two Mechanisms for Disinhibition in Identifiable Neurons in Aplysia, Physiologist, 8, 167, 1965.

84. Kornacker, K., Personal Communication, 1968.

85. Schipperheyn, J. J., Contrast Detection in the Frog's Retina, Acta Physiol. Pharmacol. Neerl., 13, 231, 1965.

86. Kuffler, S. W., Discharge Patterns and Functional Organization of Mammalian Retina, J. Neurophysiol., 16, 37, 1953.

87. Hubel, D. H., Single Unit Activity in Lateral Geniculate Body and Optic Tract of Unrestrained Cats, J. Physiol., 147, 226, 1960.

88. Hubel, D. H., Wiesel, T. N., Shape and Arrangement of Columns in Cat's Striate Cortex, J. Physiol., 165, 559, 1963.

89. Hubel, D. H., Wiesel, T. N., Receptive Fields and Functional Architecture in Two Non-Striate Visual Areas (18 and 19) of the Cat, J. Neurophysiol., 28, 229, 1965.

90. Krnjević, K., Randić, M., Straughan, D. W., An Inhibitory Process in the Cerebral Cortex, J. Physiol., 184, 16, 1966.

91. Krnjević, K., Randić, M., Straughan, D. W., Nature of a Cortical Inhibitory Process, J. Physiol., 184, 49, 1966.

92. Armstrong, C. M., Monosynaptic Activation of Pyramidal Cells in Area 18 by Optic Radiation Fibers, Exptl. Neurol., 21, 413, 1968.

93. Armstrong, C. M., The Inhibitory Path from the Lateral Geniculate Body to the Optic Cortex in the Cat, Exptl. Neurol., 21, 429, 1968.

94. Globus, A., Scheibel, A. B., Pattern and Field in Cortical Structure: The Rabbit, J. Comp. Neurol., 131, 155, 1967.

95. See refs. (94), (97), (98).

96. Dempsey, E. W., Morison, R. S., The Production of Rhythmically Recurrent Cortical Potentials After Localized Thalamic Stimulation, Am. J. Physiol., 135, 293, 1942.

97. Globus, A., Scheibel, A. B., Synaptic Loci on Visual Cortical Neurons of the Rabbit: The Specific Afferent Radiation, Exptl. Neurol., 18, 116, 1967.

98. Globus, A., Scheibel, A. B., Synaptic Loci on Parietal Cortical Neurons: Terminations of Corpus Callosum Fibers, Science, 156, 1127, 1967.

99. Polyanskii, V. B., Cycles of Excitability in Neurons of Visual Cortex in Waking Rabbits in Response to Double Flashes, Zhurnal Vysshei Nervnoi Deyatel'nosti Imeni I. P. Pavlova, 17 (4), 714, (1967). Translation from Russian in Neurosci. Transla., 4, 383, 1967-1968, Fed. Am. Soc. Biol.

100. Lettvin, J. Y., Maturana, H. R., Pitts, W. H., McCulloch, W. S., Two Remarks on the Visual System of the Frog, In Rosenblith, W. A. (Editor), "Sensory Communication", pp. 757-776, MIT Press, 1961.

101. Scheibel, M. E., Scheibel, A. B., Anatomical Basis of Attention Mechanisms in Vertebrate Brains, "The Neurosciences", pp. 602-614, Rockefeller Univ. Press, 1967.

102. Zanchetti, A., Subcortical and Cortical Mechanisms in Arousal and Emotional Behavior, In "The Neurosciences", pp. 602-614, Rockefeller Univ. Press, 1967.

103. Campbell, H. J., "Correlative Physiology of the Nervous System", p. 271, Academic Press, 1965.

104. Magoun, H. W., "The Waking Brain", Charles C. Thomas, 1958 and 1963.

105. Hernandez-Peon, R., Reticular Mechanisms of Sensory Control, In Rosenblith, W. A., (Editor), "Sensory Communication", pp. 497-520, MIT Press, 1961.

106. Eccles, J. C., Ito, M., Szentágothai, J., "The Cerebellum as a Neuronal Machine", pp. 305-306, Springer-Verlag, 1967.

107. Varbanova, A., "Interotseptivnaia Signalizatsiia (Interoceptive Signalization)", Izdatel'stvo Bolgarskoi Akademii Nauk (Publishing House of Bulgarian Academy of Sciences), Sofia, 168 pp., 1967.

108. Glees, P., "Experimental Neurology", Clarendon Press, 1961.

109. Roberts, T.D. M., "Neurophysiology of Postural Mechanisms", Plenum Press, 1967.

110. Chernetski, K. E., Sympathetic Enhancement of Peripheral Sensory Input in the Frog, J. Neurophysiol., 27, 493, 1964.

111. Cannon, W B., "The Wisdom of the Body", 2nd ed., W. W. Norton, 1939.

112. Ganong, W. F., Neuroendocrine Integrating Mechanisms, in "Neuroendocrinology", Vol. 1, pp. 1-3, Academic Press, 1966.

113. Martini, L., Ganong, W. F. (Editors), "Neuroendocrinology", Vols. I & II, Academic Press, 1966 and 1967.

114. Iberall, A., McCulloch, W. S., 1967 Behavioral Model of Man. His Chains Revealed, Currents in Modern Biol., 1, 337, 1968.

115. Robinson, B. W., Vocalization Evoked from Forebrain in Macaca Mulatta, Physiol. Behav., 2, 345, 1967.

116. Delgado, J. M. R., Evolution of Physical Control of the Brain, James Arthur Lecture on Evolution of the Human Brain, Amer. Museum Natural History, New York, 1965.

117. Pinneo, L. R., Electrical Control of Behavior by Programmed Stimulation of the Brain, Nature, 211, 705, 1966.

118. Barondes, S H., Cohen, H. D., Delayed and Sustained Effect of Acetoxycycloheximide on Memory in Mice, Proc. Nat. Acad. Sci., 58, 157, 1967.

119. Griffith, J. S., "A View of the Brain", Oxford Science Research Papers 1, Clarendon Press, 1967.

120. Schenkein, I., Levy, M., Bueker, E. D., Tokarsky, E., Nerve Growth Factor of Very High Yield and Specific Activity, Science, 159, 640, 1968.

121. Levi-Montalcini, R., Growth Control of Nerve Cells by a Protein Factor and Its Antiserum, Science, 143, 105, 1964.

122. Wertheim, G. A., Conner, R. L., Levine, S., Adrenocortical Influences on Free-Operant Avoidance Behavior, J. Exp. Anal. Behav., 10, 555, 1967.

123. DeWied, D., Opposite Effects of ACTH and Glucocorticosteroids on Extinction of Conditioned Avoidance Behavior, Proc. Int. Cont. Hormonal Steroids, 2nd, Milan 1966, 945-951, 1967.

124. DeWied, D., Inhibition Effects of ACTH and Related Peptides on Extinction of Conditioned Avoidance Behavior in Rats, Proc. Soc. Exp. Biol. Med., 122, 28, 1966.

125. Jacobson, M., Starting Points for Research in the Ontogeny of Behavior, in "Major Problems in Developmental Biology", pp. 339-383, Academic Press, 1967.

126. Hubel, D. H., Wiesel, T N., Receptive Fields of Cells in Striate Cortex of Very Young, Visually Inexperienced Kittens, J. Neurophysiol., 26, 994, 1963.

127. Wiesel, T. N., Hubel, D. H., Effects of Visual Deprivation on Morphology and Physiology of Cells in the Cat's Lateral Geniculate Body, J. Neurophysiol., 26, 978, 1963.

128. Wiesel, T. N., Hubel, D. H., Single-Cell Responses in Striate Cortex of Kittens Deprived of Vision in One Eye, J. Neurophysiol., 26, 1003, 1963.

128a. Sperry, R. W., Mechanisms of Neural Maturation, In Stevens, S. S. (Ed.), "Handbook of Experimental Psychology", pp. 236-280, John Wiley & Sons, 1951.

129. Gaze, R. M., Regeneration of the Optic Nerve in Xenopus Laevis, Quart. J. Exptl. Physiol., 44, 290, 1959.

130. Maturana, H. R., Lettvin, J. Y., McCulloch, W. S., Pitts, W H., Evidence that Cut Optic Nerve Fibers in a Frog Regenerate to Their Proper Places in the Tectum, Science, 130, 1709, 1959.

131. Gaze, R. M., Jacobson, M., Székely, G., The Retino-Tectal Projection in Xenopus with Compound Eyes, J. Physiol., 165, 484, 1963.

132. Gaze, R. M., Jacobson, M., Székely, G., On the Formation of Connections by Compound Eyes in Xenopus, J. Physiol., 176, 409, 1965.

133. Jacobson, M., Retinal Ganglion Cells: Specification of Central Connections in Larval Xenopus Laevis, Science, 155, 1106, 1967.

134. Raven, C. P., Chapters VIII and IX in "An Outline of Developmental Physiology", Pergamon Press, 3rd Revised English Edition, 1966.

135. Jacobson, A. G., The Roles of Neural and Non-Neural Tissues in Lens Induction, J. Exp. Zool., 139, 525, 1958.

136. Stone, L. S., Polarization of the Retina and Development of Vision, J. Exp. Zool., 145, 85, 1960.

137. Grafstein, B., Transport of Protein by Goldfish Optic Nerve Fibers, Science, 157, 196, 1967.

138. Sperry, R. W., The Growth of Nerve Circuits, Scientific Amer., 201, 68, Nov. 1959.

139. Arora, H. L., Sperry, R. W., Selectivity in Regeneration and Reconnection of the Oculomotor Nerve in Cichlid Fishes, Anat. Record, 148, 357, 1964.

140. DeLong, G. R., Coulombre, A J., Development of the Retinotectal Topographic Projection in the Chick Embryo, Exptl. Neurol., 13, 351, 1965.

141. DeLong, G. R., Coulombre, A. J., The Specificity of Retinotectal Connections Studied by Retinal Grafts onto the Optic Tectum in Chick Embryos, Develop. Biol., 16, 513, 1967.

142. Roberts, R G., Flexner, L. B., A Model for the Development of Retina-Cortex Connections, Amer. Scientist, 54, 174, 1966.

143. Jacobson, M., Cessation of DNA Synthesis in Retinal Ganglion Cells Correlated with the Time of Specification of Their Central Connections, Develop. Biol., 17, 219, 1968.

144. Székely, G., Untersuchung der Entwicklung optischer Reflex-Mechanismen an Amphibien Larven, Acta Physiol., Hung., 6, Suppl. 18, 1954.

145. Kornacker, K., Some Properties of the Afferent Pathway in the Frog Corneal Reflex, Exptl. Neurol., 7, 224, 1963.

146. Kornacker, K., Wall, P. D., How the Location of a Stimulus is Correctly Signalled to the Central Nervous System, 1963 Bionics Symp., "Information Processing by Living Organisms and Machines", ASD-TDR-63-946, pp. 89-92, 1964.

147. Robbins, N., Peripheral Modification of Sensory Nerve Responses after Cross-Regeneration, J. Physiol., 192, 493, 1967.

148. Jacobson, M., Baker, R. E., Neuronal Specification of Cutaneous Nerves Through Connections with Skin Grafts in the Frog, Science, 160, 543, 1968.

149. Jacobson, M., Baker, R. E., Development of Neuronal Connections with Skin Grafts in Frogs: Behavioral and Electrophysiological Studies, J. Comp. Neur., 137, 121, 1969.

150. Dubowitz, V., Cross-Innervated Mammalian Skeletal Muscle: Histochemical, Physiological and Biochemical Observations, J. Physiol., 193, 481, 1947.

151. Romanul, F C. A., VanDer Meulen, J. P., Reversal of the Enzyme Profiles of Muscle Fibres in Fast and Slow Muscles by Cross-Innervation, Nature, 212, 1369, 1966.

152. Hník, P., Jirmanová, I., Vyklický, L., Zelená, J., Fast and Slow Muscles of the Chick after Nerve Cross-Union, J. Physiol., 193, 309, 1967.

153. Hughes, A. F. W., "Aspects of Neural Ontogeny", Logos Press and Academic Press, 1968.

XI. THE CHEMICAL ORIGINS OF MAMMALIAN BEHAVIOR

Or life begins at '300'. We may regard this catch phrase as shorthand for the statement that the very simple forms of cellular life require a minimum of perhaps 300 enzyme chains to run their metabolic machinery. However, as far as the behavior of the organism is concerned, it operates with perhaps only two modes. As we climb phylogenetically toward more complex organisms, the number of enzyme chains and the number of behavioral modes both increase.

Two of the preceding sections of this report have dealt with behavior, one from a neurophysiological approach, and the other from an endocrinological-metabolic approach. In this present section we extend the discussion further, and attempt to unify the two earlier descriptions.

Human behavior can be viewed from many standpoints. For example, the psychoanalytic view emphasizes the genetic and historical origins of present behavior. As Kurt Lewin has pointed out (1), psychoanalytic theories of behavior are Aristotelian in their emphasis on historicity in causality. In sharp contrast is the Galilean view which emphasizes that present behavior is a vector summation of the forces and fields now acting.

In this section we analyze human behavior from still another point of view by asking whether the evidence previously presented showing that the chemical machinery of the living organism may be characterized by limit cycles, permits a spectral analysis of the behavior of man. Do we find, when we look at human behavior, distinguishable, recurring modes?

1. Periodic, Modal Behavior

(a) The modal view of behavior. - In the preceding sections on endocrinology, metabolism and neurophysiology, the concept that behavior is modal was introduced. However, as Kilmer, McCulloch and Blum (2) point out, it is not customary to describe human behavior in terms of modalities. The effort to do so has been recently led by Iberall and McCulloch (3), who list twenty distinguishable 'action modes'. These modes have been reconsidered and renamed in Table 1 which includes synonyms for some of them. We will not assert that these modes are fully independent, and we admit that they certainly are not 'normal modes' in the physical sense. Yet, as will be described below, the list appears to furnish an upper bound to the number of salient mammalian behavioral modalities.

(b) Separability of modes. - In our experience, different individuals of an American educated, upper middle class, when given the assignment to analyze their own behavior into separable modes, independent of content, will choose approximately twenty primary modes. After corrections are made for synonyms, the modes agreed upon are those we have shown in Table 1. Thus, this decomposition of human behavior has at least the validity that adults of our culture who are highly educated and used to introspection have little difficulty in agreeing upon it. This result was reassuring, and we take it as one indi-

cation of the reality and separate identity of the modes designated. In comparing this result with the decompositions of behavior suggested by ethologists from a number of cultures other than American middle class (however, all educated and from technologically oriented western cultures), as they describe the behavior of other mammalian species, there are differences. Nevertheless, the basic poles of behavior are not really that different.

The essential question that arises about the separate modes in Table 1 has to do with the extent to which they are mutually exclusive. Some modes occur alone, but others can be clustered and combined simultaneously with other modes. For example, it is possible to be jealous while eating, or to be sexual while relating to others, whereas it is not possible to be sexual and to be fearing, simultaneously. The possibility of simultaneity does not belie the separability of the clustered modes. Only if the modes always clustered would the separability be questionable. There are no obligatory clusters in the list of modes presented in the table.

(c) Periodicity of behavior. - Table 1 also presents a frequency spectrum for behavior which shows that individuals recognize in themselves a regular recurrence of some modes. Human memory is such that an individual can sharply recall his own behavioral states for no more than several days past, except for certain highlights ("I was irritable just before my menstrual period" or, "I had a good time on my vacation in Hawaii".) Therefore, it is not surprising that the longest period remembered for recurring behavior indicated in Table 1 is only one week. However, as will be seen later, observers may detect even longer periods in recurrent behavior.

If periodicities do exist, we can then seek the physical-chemical basis of the underlying, cyclic mechanisms and concern ourselves with the transitional probabilities among the various modes. Because of the apparent richness of human behavior it is difficult to extract a physical-chemical basis for transitional probabilities. To do so we must filter out those very details of behavior that make individuals interesting. However, if we succeed in doing so, the underlying machinery may then lie more clearly exposed.

We also can ask whether an individual personality is characterizable by a unique spectrum describing the frequency with which he enters each modal state, or we might ask whether the behavior of an individual is a Markovian process: to what extent does the path through n past states determine the likelihood of entrance to a particular future state?

Not all behavior recurs periodically. Of the twenty modes shown in Table 1, five seem to be random and intermittent. We shall question that appearance.

(d) Reactive versus endogenously-cued modal behavior (non-reactive). - Modes of behavior, whether they are randomly intermittent or periodic, are defined as 'reactive' if they occur only in response to an external stimulus. All the modes in Table 1 can be sometimes reactive, in the sense that they may occur at least occasionally as a direct response to an external environmental cue. Because recurring modes of behavior might be entrained by periodic environmental cueing, their existence does not necessarily imply also the existence

of an endogenous oscillatory phenomenon. In Table 1 we present the proposed frequency spectrum for periodic behavior, and attempt to distinguish 'reactive' (socially or environmentally entrained) rhythms from endogenous rhythms.

(e) Sex differences. - Except for the unique reproductive, maternal mode in women (gestating, mothering, nursing and nest-building) we find no sex difference in the modes accessible to men and women. Because the maternal state is uniquely female, we have omitted it from the table - especially so, since not all women necessarily enter into that state.

(f) Ergodic hypothesis applied to recurrence of non-reactive behavior. - The ergodic assumption is made concerning the modes listed. That it, it is hypothesized that within some epoch - short relative to his lifetime - every individual will manifest all the behavioral modes open to a population of humans. The maternal state violates this hypothesis, but no other mode, so far as we are yet aware, does so. With these justifications of the separable modes in Table 1, we consider the problem of behavior to be that of explaining trajectories of life as sequences of modes, and of identifying the internal cueing that displaces the human biosystem from one marginally-stable state (mode) into another.

Even superficial reflection about behavior of other animals underscores the marginal nature of the behavioral stability of animal systems: no single mode of behavior can be indefinitely sustained. But is there, for a mode of behavior, an associated time scale? If so, when and how often?

(g) Content, context and indeterminant gain. - Three propositions concerning human behavior seem to us to have such general applicability that we repeat them here, even though only the third is especially relevant to our present analysis:

1. Content of a behavioral mode is determined by both specific Aristotelian (historical, genetic) and Galilean (fields and forces now present) influences.

2. Context is the operator (in triggering reactive behavior).

3. Animal systems have indeterminate gain at zero frequency.

The content of behavior that is characteristic of an individual in a particular mode is determined by his personal history, genetics, memories, experience and psychodynamic processes. The many current theories of human behavior extract psychodynamics from such elements as genetic determination of personality, imprinting, instinct, learning, trauma, or hypothesized conflicts among unproven brain activity complexes labelled 'id, ego, superego', or 'conscious, subconscious, unconscious, preconscious', etc. Certainly, past experiences and at least three aspects of personality create casts of mood that can color behavior in many or all of our specified modes. These three aspects are: 1) anxiety; 2) depression, pessimism; 3) joy, optimism, sense of well-being. These give color to the modes of behavior but are themselves perhaps not modes. They form part of an individual's psychodynamics. We

discuss later the possibility that the coloration itself is periodic and comes in tides, as opposed to the other possibilities that it is entirely reactive, or permanently determined in advance genetically.

Our analysis does not diminish the need for a theory of psychodynamics that emphasizes specific behavioral content and its origins, for we do not explain why individuals are musicians or mathematicians, nor why they are optimists, with cups half full, or pessimists, with cups half empty. We do not ask why a person is a homosexual or heterosexual, nor why he is masturbating or fantasizing instead of being engaged in coital activity, but instead we inquire why he is in the sexual mode of behavior at all, independent of its content and particular expression. We add to psychodynamic theories an emphasis on the underlying system structure, in the hopes that coupling between the two viewpoints may be revelatory beyond present theories alone. Our theme will be that the trajectory taken through the twenty accessible modes shown in Table 1 is neither random, nor freely chosen, but is urged by internal chemical cues determined largely by thermodynamic and stability requirements. The internal cues select the responses that can be made to geophysical or social timing signals or other external perturbations which link the internal behavioral state to the outside world. In the case of some behavior the external influence is overwhelming, and in the case of other behavior it is weak.

The proposition that 'context is the operator' specifies that an observer of human behavior cannot correctly interpret individual behavior without knowledge of the context in which it occurs. We have noted the irritation with which perceptive people respond to multiphasic personality tests featuring questions such as: "When you are in a group, do you prefer to have others lead?" All but the test-makers know that this question has no answer unless it is known what the group is doing, how large it is, who are the members, and whether the individual being questioned is tired, anxious, or happily restful after some loving experience. The answers to the latter questions give the context in which the former finds an answer. A rose is not a rose is not a rose.

The proposition about indeterminate gain at zero frequency acknowledges that in a surround without fluxes or perturbations, the basic behavioral instability of animal systems will become manifested by the inevitable adoption of some behavior other than sustained rest.

(h) Identification of the present mode - the body image. - Our view of behavior does not suppose or require conscious knowledge of the present mode or the signals that lead to its adoption. Some of the signalling does reach consciousness, but far more does not. Modes themselves are not always immediately recognized by the person sustaining them. A person can be jealous and deny that fact even to himself, and insist that he is in some other mode. Consciousness, to the extent that it exists, is a narrow angle distorting lens through which we view the present. Whatever its mechanisms, we conclude that it enriches, but does not alone freely select, behavior. In the absence of the cerebral cortex, consciousness is lost. The behavioral modalities open to the decorticate animal therefore become matters of interest. To paraphrase Kilmer, McCulloch and Blum (2):

> Decerebrate cats possessing no brain above the brain stem can distinguish between tones, but not tunes. They can see brightness and attack small moving forms, but are devoid of all visual refined perception. They do not orient well to bodily touch, cold, pressure or shocks. Movements are impoverished and stereotyped; they are modal. They lie, sit, stand, walk, run, fight, surrender, eat, drink, vomit, defecate, micturate and mate. They do not pounce, cuff, or play skillfully. They are stimulus-bound and have no capacity for long-range intentions or complex problem solving. Their residual central nervous system is busy just keeping them alive. They cope only with the most urgent and important stimulus contingencies.[1]

The stimulus-bound, impoverished character of the behavior of decerebrate animals might at first appear to counter our proposition that the cortex does not alone freely select behavior. However, in our hypothesis, the cortex is always biased by the mode-selecting actions of the reticular formation and the limbic system. Models of this mode-selecting process have been attempted (2, 4). In these models, the cortex, including its property of consciousness, chooses behavior from a set of modalities being urged upon it by the reticular system of the brain. The reticular system determines the priorities, and these change with time. Therefore, for certain periods the cortex can override an urging from the reticular core, and reject a particular mode being urged, but ultimately the cortex may be driven to make the choice of that mode by increasing stimulation from the reticular system. Thus, the extent to which the cortex 'freely chooses' behavior varies according to the strength of the insistence of the reticular system that a particular mode be chosen. A three brain view of man stressing these compartments and the limbic system has been strongly sponsored by MacLean.

Following Kilmer and McCulloch, we postulate that some image of the present state or mode, and the transitional probabilities for change of mode are computed by parts of the brain other than the cortex - presumably by the reticular formation and the limbic system. From the earlier discussion of the neurophysiological aspects of mind, we can consider what might form the basis of the computation. The brain has three possible paths of information about the present state of the body: 1) the special senses (seeing, hearing, smelling, tasting, surface reporting - thermal, pain, touch receptor reports from skin and mucous membranes); 2) visceral neural sensations (fullness, pain, nausea, proprioception, muscle tension, dizziness, impending syncope); and 3) the levels or rates of change of levels of chemicals - hormones, drugs, metabolites - in blood, providing these substances are capable of affecting the above two pathways, or of crossing the blood-brain 'barrier'.

[1] On the other hand, preliminary studies of Katcher, Turner, Jacobowitz (personal communication), with puppies whose peripheral sympathetic nervous systems have been abolished (by 6-hydroxydopamine), show little evidence of any but 'normal' gross behavior.

We believe the above list of sensory reports includes most, and perhaps all possible reports, unless special states such as impending suffocation - as occur in drowning or pulmonary edema - are unique, and not merely combinations of other reports. (Some psychiatrists have referred to "the image of the heart".)

With respect to the creation and updating of a body image by use of the special or visceral sensations, we suppose that the signalling is straightforward, that the pathways are defined anatomically, and that local receptor characteristics determine whether the receptor is activated by stretch, by temperature change, by potassium leakage from nearby cells, etc. Therefore, our emphasis instead is on the problem of how the brain creates and updates a body image based upon the profile of chemicals in blood that can cross the blood-brain barrier in their original or transformed state.

(i) The chemical basis of a body image. - The brain could, in principle, read hormones, drugs or metabolites in blood providing these can cross the blood-brain barrier. We define metabolites as substances, or their biotransformation products that serve as fueld or building blocks for tissue growth, maintenance, or activity. In the main, these will be sugars, amino acids, fatty acids or their derivatives, or cofactors such as vitamins and minerals. There is, however, another class of chemical signals in blood that arise from metabolic activity or immunoreactivity of peripheral cells and tissues. We designate this class of substances as 'autocoids' (5). The autocoids include: histamine, bradykinen, kallikrein, angiotensin, serotonin. Many of these chemicals are listed in Table 2.

It is important to note that the neural transmitters in the peripheral nervous system and spinal cord, acetylcholine, GABA, norepinephrine and glycine cannot be directly read from blood into brain because they do not readily pass the blood-brain barrier, except perhaps glycine. However, a representation of the body state might consist of the distribution of central nervous system neural transmitters within the brain. Five substances are known which could play this role within the brain: acetylcholine, norepinephrine, serotonin, GABA and dopamine. Substance P may also have a central neurotransmitter function. If the body image is represented by the distribution of transmitters in the brain, which may in turn reflect prior patterns of electrical activity and synaptic growth, then a false image - a phantom - might be produced were the brain itself to suffer a derangement in the metabolism of its own synaptic transmitters. An aberrant trajectory through the modes of behavior might then result. The content of the modes might also be profoundly altered.

(j) Updating the body image. - Whether the body image is continuously created or intermittently updated is not known.

2. Chemical Communication Pathways To and From Brain

The earlier section in this report dealing with the brain introduced the subject of neuroendocrinology. Here we apply the information more directly to human behavior.

(a) The blood-brain barrier. - The blood-brain barrier is both an anatomical (6) and functional barrier (7). Anatomically it consists of a molecular sieve: the capillary endothelium, basememt membrane and glial cell foot processes. Functionally it includes enzymes in or on the endothelial cells that alter substances passing through the capillary barrier, as well as selective active transport processes and pinocytosis. The barrier, being functional, is not constant. An increase in neuronal activity or thyroid function increases the uptake of labelled amino acids into neurons (8). In fact, with respect to most amino acids, the term 'barrier' is somewhat inappropriate, yet GABA is excluded from brain (7). For each substance it must be separately determined whether transfer from blood to brain can occur readily or not.

(b) Chemical communication pathways from brain - I. Neurosecretion. - Just as a chemical signal flow from blood to brain is selective, so is that from brain to blood. Normally, as far as is now known, the brain delivers only two chemical signals into the general circulation (other than the usual metabolites such as CO_2 resulting from its own respiration and metabolism). These two substances are the posterior pituitary hormones oxytocin and vasopressin. The delivery process is called 'neurosecretion'.

(c) Chemical communication pathways from brain - II. Releasing/inhibiting factors affecting the adenohypophysics. - The brain delivers at least seven chemical signals to the pars tuberalis, distalis or intermedia of the pituitary through the pituitary-portal circulation (9). These substances have been mentioned in the preceding section of this report dealing with neurophysiological principles. They are the hypothalamic releasing and inhibiting factors and they include: follicle stimulating hormone releasing factor (FRF), luteinizing hormone releasing factor (LRF), prolactin inhibiting factor (PIF), corticotropin releasing factor (CRF), melanocyte stimulating hormone releasing factor (MSH-RF), thyroid stimulating hormone releasing factor (TRF), and growth hormone releasing factor (GRF). A releasing factor for prolactin (PRF) and an inhibiting factor for growth hormone (GIF) may exist also. Although all these substances are stored in the medial, basal hypothalamus, they do not necessarily arise there. The evidence is not yet convincing, but a pathway from neurons or glia through cerebral spinal fluid to the ependymal cells of the third ventricle and thence to the pituitary portal system may exist, and it could provide chemical communication between many parts of the brain - other than hypothalamus - and the adenohypophysis. The evidence for this pathway has recently been critically reviewed (9).

The anterior pituitary in any case serves as an amplification stage in chemical transmission from brain to the endocrine systems involving the thyroid, gonads, adrenal glucocorticoid cortex, growth hormone, prolactin and MSH; these systems affect the state of much of the rest of the body.

Much more is known about the chemical transmission from brain than about chemical transmission to brain. Both processes occur, and are essential in the analysis of behavior we develop here, because the parameters of the metabolic machinery that determine the marginally-stable, oscillatory state adopted or set by hormones as well as by nerves.

Examples of chemical transmission to brain will be given later, since this aspect of neuroendocrinology is new and rapidly-developing. The older branch of neuroendocrinology, that of chemical transmission from brain through the adenohypophysis to the rest of the body will not be detailed, since it is now a mature and extensive subject, rather well developed, as noted in a preceding section of this report. A recent review concerns this material (9).

3. Thermodynamics and Behavior

(a) Summational invariances. - Closed statistical mechanical systems are characterized by at least three summational invariants: 1) conservation of momentum; 2) conservation of mass; 3) conservation of energy.

In the case of biosystems which are thermodynamically open, exchanging both mass and energy with their surroundings, other summational invariances may be present. If these could be identified, much of the design and behavior of complex biological systems might seem simplified and more easily understood. Unfortunately, the additional summational invariants are not as yet identified, and so far as we know, the matter of their existence has not previously been as explicitly raised as we do here. We suspect, however, that the 'value function' proposed by Manfred Eigen, discussed in an earlier section of this report, may involve an unidentified invariance of this kind at the molecular level.

(b) Closure over duty cycles. - For the present, we can only point out again what the biologist has long known - namely, that over any duty cycle, a biosystem must meet the thermodynamic requirements of energy and mass balance if it is ever to recover its initial condition and preserve its identity. In periods short relative to his lifetime, a human is assumed to recover to a close approximation any particular initial condition repeatedly, even though we recognize that in the time domain of seventy years he is aging progressively and irreversibly so that exact recovery of a prior state is impossible.

(c) Behavioral modes most clearly related to thermodynamic requirements. - It is to be expected that these thermodynamic issues are assured of settlement by appropriate behavior strongly driven. The first five items in Table 1 show examples of strongly-driven, recurrent behavior modes, all concerned with recovery of the thermodynamic initial condition after a duty cycle consisting of some trajectory to other modes of behavior. Resting, eating, drinking, sleeping and voiding are unmistakably strongly obligated behavioral modes, not easily overriden even temporarily by 'will'.

(d) Internal behavioral cues. - The internal cues for eating, drinking, or voiding are well identified in physiological literature. They involve

receptors and neural signals sensitive to physical (stretch, muscle tension in stomach walls) and chemical (blood levels of glucose, and amino acids) or physical-chemical (plasma or cerebrospinal fluid osmolality) variables, or their rates of change. The stretch and tension receptors are in the peripheral nervous system, but the chemical detectors are in the brain itself, as noted in an earlier section. Reports from these receptors ultimately cause the brain (the reticular formation and limbic system) to commit the cortex to a mode of behavior suitable for reducing the receptor nerve impulse traffic. The cyclic behavior resulting is relaxational in character, and the analysis is not particularly difficult. Periodic mass intakes (food, water) and outputs (urine, feces) result.

The internal cues for rest and sleep have not been so well identified. However, the dietary and motor activity patterns of humans both tend to produce acidosis, and so we propose that at a time at which the profile of chemical concentrations defining the acidosis has reached some critical point in the brain, the rest mode is switched on. There is a slow respiratory blow off and renal excretion of acid products with a time constant of the order of four hours. In four hours, one is sufficiently rested so chat he could resume an activity state. However, for the circadian, sleep-wakefulness cycle signal to clearly turn on, the relaxation must be more complete - and extended for at least a few time constants of the acidosis-reducing process. This extension comprises the nearly eight-hour rest period customary for sleep. Thus, the rest state is viewed as being a follower, and is not itself a timing signal.

In the case of eating, the reinforcing daily buildup of somnambulance - small after breakfast, bigger after lunch, almost irresistible after dinner (indicating insipient instability) - is quite apparent.

It is possible that the acidic condition puts the nervous system into an unstable state between two brain-states, characterized during sleep as REM and non-REM. (These two states are possibly identifiable during wakefulness also, we believe). During this unstable state the dreaming (REM) instability emerges, whose experimental occurrence provides evidence of the biological need for a fixed number of relaxing phases, occurring as relaxation oscillations. Thus dreams may be involved with the acidic biochemical state of the brain. It isn't until the acidotic processes and byproducts are sufficiently reduced that the circadian clock will turn the activity pattern on again. All of these opinions are, of course, quite speculative, but they fit known facts together, and suggest models that can be experimentally probed.

With regard to the other periodic, non-reactive behavior shown in Table 1, the analysis is handicapped by lack of knowledge concerning the internal cues. At least seven modes listed fall into this category: grooming, changing posture, using body in motor play, working, being sexual, relating to other, fantasizing. Of these modes, the first four involve motor states, and three (being sexual, relating to others, fantasizing) do not originate as motor states.

(e) Modes without assigned periodicities. - The residue of modes in Table 1 all involve 'states of mind' that may ultimately manifest themselves

in motor actions, but need not do so initially, and the individual observer is not able to detect periodicity in these modes in himself, nor are the internal cues known.

4. Long-Term Periodicities of Human Behavior

Before proceeding further in the analysis of human behavior, it is necessary to decide whether all the periodicities actually present are designated in Table 1, which was based upon self-analysis of behavior. Richter (10, 11) has presented strong evidence that normal humans have 'clocks' with periods of 1½-2 hours, 24 hours, 28-30 days, 280 days. These are shown in Table 3A, modified from Richter. From the study of one individual with periodic somatic or mental illnesses, striking evidence for the existence of 'clocks' with periods from 12 hours to 2 years has been obtained (Table 3B). In one case, a 10 year period of recurrence was observed, but with such a long period, very few cycles can be observed in a lifetime, so that the proof of true periodicity is difficult.

We will not here describe the diseases that manifested the indicated periods, since Richter has done so in detail (10). We wish to emphasize though, his conclusion that the diseases did not produce the periods: the presence of the disease merely provided an unaccustomed, conspicuous marker or 'hand' to reveal the presence of an underlying (presumably normal) clock mechanism. While this conclusion can be challenged, we tentatively accept it, and then turn our attention to the nature of the underlying clocks.

The existence of cycles with long periods raises the possibility that many or all of the behavioral modes listed in Table 1 as non-periodic, may in fact be periodic over intervals too long to be remembered during self-reflection. New observations are very badly needed to confirm or deny the contention that all the modes of behavior listed in Table 1 are periodic. Furthermore, even the global colorations of modes (anxiety, depression, joy) previously described, could be similarly periodic.

(a) The modal hypothesis extended. - The facts are not yet clear, but we predict that the clocks with long periods described by Richter are present in normal humans and they they manifest themselves through periodicities in the 20 modes and 3 colorations of modes that we believe encompass most human behavior. Thus, in summary, we hypothesize that all non-reactive behavior is both modal and periodic.

(b) Timers and followers. - Some behavior that is periodic, modal and non-reactive occurs following limit-cycles in metabolic processes associated with other modes. The occurrence of rest or sleep following eating and activity was given as an example above. In some cases it is not easy to discern which modes and processes are the followers, and which are the timing oscillations entraining them. A follower may serve as an intermediate stage between a timing oscillator and another follower.

Thus, for example, the circadian activity timer, whatever it may be, entrains such modes as rest; eating, drinking; mineral balances; activity; or daily psychic changes. Each in turn, acting as a subinterval timer, may cue another of these modes as a subsequent follower. Longer-term timers would include the menstrual cycle in women, for example.

5. The Full Spectrum for the Life Phase of Adults

(a) The zones of the spectrum. - The longest period in a man's life is his life itself, from fertilization of the ovum to death. As with the life of any factory, the total life of man may be subdivided into the stages of construction (growth, development, differentiation), start-up, regulated maturity and decline. Although the decline of man begins almost immediately after cessation of growth, it is sufficiently slow in most cases to justify our designating the ages of 18 to 60 as regulated maturity.

(b) Frequency ranges in regulated maturity. - During regulated maturity the various timers and entrained follower processes oscillate with periods ranging from milliseconds to several years. The line-spectrum that we believe describes the human biosystem can be conveniently subdivided into high frequency (0.1-100 sec), medium frequency (2 minutes to one day) and low frequency (one day to two years) ranges. Some of the medium and low frequency processes and behavior have already been discussed (Tables 1 and 3). Earlier, provisional accounts of the spectrum have been published (13, 14).

In addition to the behavioral spectra suggested by Tables 1 and 3, it would be desirable to have a spectrum for the individual metabolic and hormonal components of the human biosystem. Specifically, for each item in Table 2 we would like to know the answers to the following questions:

1. What is the lag time to the first detectable effect when the substance is given (from several initial conditions or operating points of the system, and by several input wave forms - since the system is nonlinear)?

2. What is the time to peak effect?

3. How fast does the signal decay when the input is removed?

4. Does the component substance reveal cyclic change in its levels when the system is not forced by it?

5. What is the period of any observed cycle?

Such a spectrum is very difficult to construct because the existing biological literature deals mainly with identification of substances and their causal relationships, but now with dynamics. Time is almost never explicitly considered in biological work, and the physically-oriented scientist looks with dismay upon the vast and variegated monument to lost opportunity for dynamic analysis enshrined in the literature of the fields of endocrinology

and metabolism. If our efforts here do no more than provoke biologists to remedy this glaring defect in their data bank, we will have served them well.

Table 4 presents a mere shadow of the component spectrum that is required for dynamic analysis. Almost no data were available for the human biosystem. Table 5 is based upon a search of the recent endocrine literature, and presents some data on time domains of hormone action in humans, but a much more extensive search is needed. We expect that entirely fresh research will have to be undertaken.

(c) Correlation of spectra. - In the section of Metabolic Systems we pointed out that it is difficult to prove causal relationships in the biosystem by means of spectroscopy, because frequency multiplications or demultiplications may occur in nonlinear systems. We added, however, that the spectroscopic approach might suggest the existence of causal relationships previously unsuspected or underemphasized.

In this section we have presented three partial spectra: a) human behavior; b) metabolic component levels in blood, and 3) hormonal levels and times of action. These three spectra can be correlated by imagining a matrix whose columns represent chemical materials, chemical process change, hormones and behavioral modes, and whose rows represent discrete time intervals (periods) from shortest to longest, encompassing all the lines found in the various spectra. Not all the intersections of rows and columns are filled. However, in any column, the many temporal levels of regulation of that material, process or mode are catalogued. In any row, the elements may or may not be coupled, but it is here that causalities (couplings) are suggested.

We have not explicitly presented the resulting matrix here, because the data are too scanty, but the principle remains important. The discussion that follows suggests some of the features that emerge even from the very incomplete matrix, and these are summarized below.

Hormone	Process or Component	Line Period
Insulin	Glucose levels	30-60 seconds
Epinephrine	Heat production Ventilation Red cell flow Glucose levels Postural changes	2 minutes
Vasopressin	Ventilation Oxygen balance Surface temperature Glucose levels Grooming	7 minutes

Hormone	Process or Component	Line Period
Vasopressin Epinephrine Corticosterone Insulin Glucagon	CO_2 balance Glucose levels	20-30 minutes
Cortisol Growth hormone	REM effect Motor activity Fantasizing	90 minutes
Cortisol MSH Growth hormone	Thermal balance Food intake Fatty acid levels Working Relating to others	3-4 hours
Cortisol Testosterone in males	Sleep-wakefulness Water and mineral balance (coarse tuning)	1 day
Aldosterone Testosterone Estradiol	Water balance (fine tuning) Mineral balance (fine tuning) Sexual feelings	3 days
Thyroxin Testosterone Growth hormone	Work-rest cycle	1 week
Progesterone Estradiol Gonadotropins	Menstrual cycle Mood cycles	1 month

Again, we repeat, we do not assert that sharing of spectral lines necessarily indicates the presence of coupling, but it does suggest the possibility. In the table given above, the couplings may be between hormones and processes or components, or among the processes and components themselves listed in the center.

6. Hormones and Behavior - Sexual Behavior

The most thoroughly studied relationship between hormones and behavior concerns sexual behavior. Effects of hormones on many aspects of behavior including sexual have recently been reviewed (19). The dependence of sexual behavior upon gonadal hormones has been obvious for centuries. For example, Leonardo da Vinci (1452-1519) applied his formidable intelligence to this matter, and in his journals he notes that he should "show how the ardour and vigour of the animals is caused by the testicles" (20). (Leonardo designated ovaries as female testicles, and so his word "testicles" is equivalent to our word "gonads".) Subsequent work has shown that Leonardo was quite correct, and both sexual and aggressive behavior in subhuman primates is dependent in part on secretory products of the gonads. It is to be expected that this dependence would have been thoroughly explored by biologists, and the literature is large.

The dependence of behavior on the gonads occurs in stages. As noted in an earlier section of this report, androgens (which are present in both sexes, with testosterone as archetype) affect brain structure if they are present in high levels during a critical period in the prenatal-neonatal period. They cause loss of the cyclic mechanism for gonadotropin release that in adult females leads to estrus or menstrual cycles. The loss is permanent. Later, in adult life, androgens affect aggressiveness and dominance (21, 22).

The major sex hormones are estradiol, progesterone and testosterone. All three substances readily cross the blood-brain barrier. The extensive recent work upon distribution of estrogens in brain has been reviewed (9). In brief, estradiol is taken up and concentrated in those regions of the brain known on the basis of studies with lesions and electrical stimulation to affect sexual behavior.

The coupling between human behavior and the menstrual cycle is well known (23, 24), but the evidence is largely anecdotal. The best-known study of human sexuality and the menstrual cycle is that of Udry and Morris (25). They found that both rates of intercourse, and achievement of orgasm were higher at the middle (ovulatory) region of the cycle, and then again just before menstruation, than at other times. Similar results on copulatory frequencies, showing a midcycle peak, have been found in the Rhesus monkey (26, 27).

In the case of monkeys, male sexual behavior cycles also (27), but it is cued by pheromones ('smells') released by the genital area of the cycling female. These smells both excite and inhibit (at different times) the male behavior. Thus, in this time domain we find strong external cueing for some of the behavior observed, but autonomous, endogenous rhythms for other components of behavior.

The androgens of the female adrenal support sexual behavior (28, 29), as does estrogen (29, 30, 31). Progesterone in the female strongly suppresses male interest in her (32, 33). It is clear that sex hormones affect subhuman primate sexual behavior, and do so in a somewhat selective manner.

The results of studies on monkeys, though striking, do not justify an uncritical application to the explanation of human sexual behavior, since much behavior in humans is conditioned or learned. However, what little evidence there is suggests that women are internally cued in their sexual behavior as are monkeys, and that periodicities relating to the menstrual cycle are common.

7. Hormones, Metabolites and Behavior - Summary

The number of clearly definable metabolic states appears to be small, as noted in the section on metabolic systems, and much less than the 20 modes shown in Table 1. However, the hormones considered as affecting the metabolic state numbered only six, whereas humans have about 40 hormones overall, and many metabolites. Since a human possesses at least 10^5 different protein species, the number of degrees of freedom for internal states could possibly be very large.

Physical scientists might think of this problem in the following way. In order to get at a continuum or macroscopic physics from an atomistic base, one tends to classify molecular-atomistic degrees of freedom as 3 translational plus 3 rotational rigid body degrees of freedom, and then other rotational and vibrational degrees of freedom are included. In the biological system, the atomistic degree of freedom includes a much larger number. The atomistic elements are made up of even more primitive elements. There are perhaps 10^3-10^4 degrees of freedom involving biochemical chains. Most of these are the protein syntheses dealing with immunological problems and self-recognition of protein by protein, which provide specificity for the organism. There is no guarantee that these 'degrees of freedom' are independent, or independently invoked.

The degrees of freedom associated with biochemical chains represent the possible maximum number of chemical modalities in the system. The great number of these possible modalities creates a tremendous amount of biological diversity. It may help to contrast this with a hydrodynamic or elastic system, which can have an infinite number of degrees of freedom (e.g., those arising from harmonic quantization by the walls), but only a few modalities. For example, the hydrodynamic field may have only a wave-like, a viscous-diffusive, or a thermal diffusive mode. Thus, most hydrodynamic systems look similar in the macroscopic domain. Biological systems, in contrast with finite numbers of degrees of freedom, but rich modalities, look diversified. However, most of the diversity is in appearance, and not in function (the structural envelope looks different; the functional action may remain the same from one species to another).

Modalities are concerned with abstract properties in space and time. If the 'atoms' have all these complex modalities, what sort of statistical mechanics are we to expect? What sort of phenomenological macro-continuum will be revealed? We do well to recognize that most of the modalities concern structural rather than functional elements, as previously mentioned, and in most operative domains they are frozen out degrees of freedom. Since components of biological systems usually show rapid turnover, in which the atomistic time domain is not far removed from the macroscopic time domain, so that the

system is 'fluid', the problem can be viewed as transformation of form into function going on at all levels at all times. In that case, it is probable that, as we have proposed, a 'quantum' spectroscopy is the appropriate analytic tool.

At the present time, in the absence of further information, we can only propose, for further investigation, that the 20 modes of behavior in Table 1 are expressions of as many constellations of hormonal-metabolic states, each of which is only marginally stable.

8. Summary of the Physical View of Behavior

The biological literature will refer the reader to homeostasis, the concept developed by Bernard, Sechenov, Cannon, by which the complex system retains a near-constancy of its internal parameters, independent of conditions in the external environment after disturbances are removed. In physical terminology, quasi-static regulation is implied, not simply a force-balancing equilibrium.

However, measurement of many physiological parameters, under steady state or very slowly changing conditions, reveals not a near-constancy or small fluctuation of simple static regulation or linear feedback control, but large scale oscillations. It is the average values which seem to be regulated, thus the term 'homeokinesis' is introduced to represent a dynamic regulation - much more nearly like bang-bang control - by which the average state of the system is achieved.

However, it is not only the internal organ systems which operate in this mode of biochemical regulation via a large number of spectral lines, but also the behavioral system (representing a generalized bio-spectroscopy).

The complex organism as a whole is not stable. If you put it down, it will soon begin to move. If wildly unstable, and placed in a field or enclosure, it will settle down for a time. Thus, the system is marginally unstable, and representable as a self-actuated motor system that intermittently hurls itself into search modes. Its primary algorithm is that it eats and moves about so that it can continue to eat and move about; and aperiodically to couple so as to produce a replicated offspring that will grow to full scale and continue to eat and move about.

Operationally, there are about 20 focal centers - hungers' - among which the system patterns its behavior. The system motion consists of open relaxational 'motional' phases, followed by locked-up orbits in which the system circles its 'foci' (which contribute to its behavioral spectrum). A successful individual is one who patterns his behavior satisfactorily among all his hungers without saturating, or becoming unstable among only a few poles. A successful species can, as a result of its genetic coding, epigenetically unfold satisfactory adaptive behavioral patterns, as the ecological environment changes, and a new higher level of behavior emerges.

Table 1

Behavior Modes of Man

Mode	Periodic?	Estimated free-running period or frequency	Entrainable by environmental (social) cues?
Resting	yes	10 min, 2-3 hrs, 5 days, 3 months	no
Eating	yes	2-4/day	yes
Drinking	yes	4 hours	yes
Sleeping REM non-REM	yes	1-2/day	yes
Voiding	yes	4/day urine 1/day feces	yes (weakly)
Grooming	yes	10 mins	yes
Changing posture	yes	1-2 mins	no
Using body (exercise, play, gross motion)	yes	90 mins	no
Working	yes	2-3/day	yes
Being sexual	yes	3 days	yes
Relating to others Loving Caring Cooperating Stroking, touching Conversing Sheltering	yes	4/day	yes
Fantasizing	yes	10/day	no
Withdrawing, escaping	no	--	--
Attending Arranging Planning Problem-solving Learning Studying	yes	--	yes

Table 1 (Cont'd.)

Mode	Periodic?	Estimated free-running period or frequency	Entrainable by environmental (social) cues?
Creating Being introspective Reading Thinking			
Being aggressive Competing Striving	yes	--	no
Contending Fighting Hating Being hostile Being angry	no	--	yes
Being acquisitive Being greedy Stealing Cheating	yes	--	no
Envying Being jealous	no	--	--
Feeling loss Grieving	no	--	--
Fearing	no	--	--

Table 2

Endogenous Substances Normally in Blood that Might Strongly Affect Mammalian Behavior

I. Hormones

A. ACTH
B. MSH
C. GH
D. AVP, LVP
E. Prolactin
F. Corticosterone, cortisol
G. T_4
H. Epinephrine, norepinephrine
I. Estradiol
J. Progesterone
K. Testosterone

II. Metabolites

A. Glucose
B. Amino acids
C. FFA
D. Cyclic AMP
E. Hydrogen ion
F. CO_2
G. Sodium ion

III. Autocoids

A. Histamine
B. Bradykinin
C. Kallikrein
D. Angiotensin
E. Serotonin

Table 3

A. Periods of behavior in normal humans

Period	Process or Mode
1½-2 hours*	General motor activity
24 hours	Sleep-wakefulness cycle, defecation
28-30 days	Menstrual cycle with attendant mood changes
280 days	Duration of pregnancy, with attendant behavioral changes

*Recently rediscovered (12)

B. 'Clock' found in patients with periodic somatic or mental illnesses

Periods observed
12 hours
24 hours
48 hours
5 days
7 days
13 days
17-19 days
21 days
24-25 days
20-30 days
40 days
40-60 days
4-5 months
1 year
1¼ years
2 years
(10 years)

Table 4

Periods or Frequencies of Cycles for Metabolic Constituents

Common metabolic constituents of arterial blood	Species	Periods or frequency of oscillations	Reference
Oxygen	guinea pig, cat, man	30-60 secs; 100-200 secs; 400-500 secs	(16,34,35)
CO_2	guinea pig, man	1-2 mins; 6-8 mins; 30-60 mins; 100 mins	(16,35)
Glucose	rats, dogs, man	40 secs; 100 secs	(15,17,18)
Fatty acids (at least 3 individual substances)		None yet looked for	
Amino acids (20 individual substances)		"	
Glycerol		"	
Lactate		"	
H^+		"	
Etc.		"	
Temp. of arterial blood	guinea pig, man	3-5 secs; 30 secs; 60 secs; 100-200 secs; 325-400 secs	(14,16)

(A summary of various data may be found in (36).

Table 5

Time Domains for Hormones

Hormone	Time to first detectable effect	Relaxation Half-life (humans)	Periods
Cortisol, Corticosterone	5 min (negative feedback)	90 min (cortisol)	3 hrs
	1-2 hrs	30 min (corticosterone)	24 hrs
	1 day		
	3 days		
Epinephrine	30 sec	1 min	
	5-15 min	1 min	
T_4, T_3	1 min	7 days (T_4)	
	10 min	1½ days (T_3)	
	6 hrs		
	1 day		
	30 days		
Estradiol	1-10 min	70 min	
	1 hr		
	6 hrs		
	12 hrs		
	48 hrs		
Testosterone	1 day	90 mins	circadian in men
	2 days		
	7 days		
Progesterone			
ACTH	2 mins	5-20 mins	
βMSH	4 hrs		
GH	20 mins	20-35 mins	
	1 hr		
	6 hrs		
	24 hrs		
	48 hrs		
	1 week		
AVP, LVP	5 min		
	15-30 min		
Prolactin			
Epinephrine	5 secs		
	5-15 mins		

Bibliography

1. Lewin, K., The Conflict Between Aristotelian and Galilean Modes of Thought in Contemporary Psychology, In: "A Dynamic Theory of Personality: Selected Papers", pp. 1-42, McGraw-Hill, New York, 1935.

2. Kilmer, W., McCulloch, W., Blum, J., Embodiment of a Plastic Concept of the Reticular Formation, In: "Biocybernetics of the Central Nervous System", pp. 213-260, Little, Brown & Co., Boston, 1968.

3. Iberall, A., McCulloch, W., 1967 Behavioral Model of Man - His Chains Revealed, NASA Contractor Report CR-858, Wash. D.C., July 1967.

4. Kilmer, W., McLardy, T., A Circuit Model of the Hippocampus of the Brain, Interim Sci. Rep. No. 8, July 30, 1970. Proj. No. 976902, Contract No. AFOSR-1023-67B, Air Force Off. of Sci. Res., Dir. of Info. Sci., Arlington, Va., 22209.

5. Goodman, L., Gilman, A., "Pharmacological Basis of Therapeutics", MacMillan, New York, 1965.

6. Quadbeck, G., Clinical Importance of Alterations in the Blood-Brain Barrier, Prog. Brain Res., 29, 343, 1968.

7. Van Gelder, N., A Possible Barrier for Gamma-Aminobutyric Acid in the Central Nervous System, Prog. Brain Res., 29, 259, 1968.

8. Ford, D., Changes in Brain Accumulation of Amino Acids and Adenine Associated with Changes in Physiologic State, Prog. Brain Res., 29, 401, 1968.

9. Yates, F., Russell, S., Maran, J., Brain-adenohypophysial Communication in Mammals, Ann. Rev. of Physiology, 33, 1971 (in press)

10. Richter, C., "Biological Clocks in Medicine and Psychiatry", Charles C. Thomas, Springfield, 1965.

11. Richter, C., Endocrine Function, Periodicity and Behavior, In: Michael, R. (Ed.), "Endocrinology and Human Behavior", Oxford U. Press, London, pp. 239-255, 1968.

12. Oswald, I., Merrington, J., Cyclical "On Demand" Oral Intake by Adults, Nature, 225, 959, 1970.

13. Iberall, A., New Thoughts in Biocontrol, In: Waddington, C. (Ed.), "Towards a Theoretical Biology. 2. Sketches", Aldine, Chicago, 1969.

14. Iberall, A., Cardon, S., Further Study of the Dynamic Systems Response of Some Internal Human Systems, NASA Contractors Report CR-219, May 1965, Wash. D. C.

15. Weinberg, M., Iberall, A., The Oscillatory Nature of Glucose Regulation (abstract), Fed. Proc., 29, 947 Abs, 1970.

16. Iberall, A., Ehrenberg, M., Cardon, S., General Dynamics of the Physical-Chemical Systems in Mammals, Sixth report, Contract No. NASW-1066, August 1966, Nat. Aero. and Space Admin., Wash. D. C.

17. Anderson, G. et al, Postabsorptive Undulations and Oscillations in Blood Glucose, Amer. J. Clin. Nutrition, 4, 673, 1956.

Anderson, G., Kologlu, Y., Papadopoulos, C., Fluctuations in Postabsorptive Blood Glucose in Relation to Insulin Release, Metabolism, 16, 586, 1967.

18. Hansen, K., Oscillations in the Blood Sugar of Fasting Normal Persons, Acta Med. Scand., Suppl. 4, 27, 1923.

19. Relkin, R., Effect of Endocrines on Central Nervous System, Part I, N. Y. J. Med., 69, 2133, 1969.

20. Belt, E., Leonardo the Florentine (1452-1519), Investigative Urology, 3, 99, 1965.

21. Edwards, D., Early Androgen Stimulation and Aggressive Behavior in Male and Female Mice, Physiol. and Behavior, 4, 333, 1969.

22. Work, M., Grossen, N., Rogers, H., Role of Habit and Androgen Level in Food-Seeking Dominance Among Rats, J. Comp. Physiol. Psych., 69, 601, 1969.

23. Janowsky, D., Gorney, R., Mandell, A., The Menstrual Cycle: Psychiatric and Ovarian-Adrenocortical Hormone Correlates: Case Study and Literature Review, Arch. Gen. Psychiat., 17, 459, 1967.

24. Janowsky, D., Gorney, R., Kelley, B., The Curse-Vicissitudes and Variations of the Female Fertility Cycle, Part I. Psychiatric Aspects, Psychosomatics, VII, 242, 1966.

25. Udry, J., Morris, N., Distribution of Coitus in the Menstrual Cycle, Nature, 220, 593, 1968.

26. Michael, R., Zumpe, D., Rhythmic Changes in the Copulatory Frequency of Rhesus Monkeys (Macaca Mulatti) in Relation to the Menstrual Cycle and a Comparison with the Human Cycle, J. Repro. Fertil., 21, 199, 1970.

27. Michael, R., Welegalls, J., Ovarian Hormones and the Sexual Behavior of the Female Rhesus Monkey (Macaca Mulatta) under Laboratory Conditions, J. Endocr., 41, 407, 1968.

28. Everitt, B., Herbert, J., Adrenal Glands and Sexual Receptivity in Female Rhesus Monkeys, Nature (London), 222, 1065, 1969.

29. Trimble, M., Herbert, J., The Effect of Testosterone or Estradiol Upon the Sexual and Associated Behavior of the Adult Female Rhesus Monkey, J. Endocr., 42, 171, 1968.

30. Michael, R., Saayman, G., Differential Effects on Behavior of the Succutaneous and Intravaginal Administration of Estrogen in the Rhesus Monkey (Macaca Mulatta), J. Endocr., 41, 231, 1968.

31. Vandenbergh, J., Endocrine Coordination in Monkeys: Male Sexual Responses to the Female, Physiol. Behav., 4, 261, 1969.

32. Michael, R., Saayman, G., Zumpe, D., The Suppression of Mounting Behavior and Ejaculation in Male Rhesus Monkeys (Macaca Mulatta) by Administration of Progesterone to their Female Partners, J. Endocr., 41, 421, 1968.

33. Michael, R., Plant, T., Loss of Potency in Male Rhesus Monkeys when Paired with Females Receiving Contraceptive Steroids, (An abstract in the Proceedings), J. Endoc., 45, i, 1969.

34. Whalen, W., Nair, P., Skeletal Muscle pO_2: Effect of Inhaled and Topically Applied O_2 and CO_2, Am. J. Physiol., 218, 973, 1970.

35. Lenfant, C., Time-Dependent Variations of Pulmonary Gas Exchange in Normal Man at Rest, J. Appl. Physiol., 22, 675, 1967.

36. Cardon, S., Iberall, A., Oscillations in Biological Systems, Curr. Mod. Bio., 3, 237, 1970.

XII. SUMMARY, OVERVIEW, AND DISCUSSION

A set of principles have been set forth by which the complex biosystem may be regarded within both a physical and biological context. With a deceptive simplicity, its paradigm might be summarized by the statement that to the spatial tool par excellence of the biologist, microscopy, we have added the temporal tool of the physicist, spectroscopy, to aid in the search of form and function.

These principles include:

Homeokinesis - Internal systems within the biological organism are caught up in closed, repetitive chains, oscillators, and the system is governed both chemically and electrically by mediating the stability and operating point of these periodically repetitive chains of events. The total collection of these oscillator chains may be regarded as making up a constellation. The mechanisms for mediating the operating points of these chains may themselves form an autonomous oscillator chain (like a ring oscillator), or they may aperiodically change in accordance with conditions in the surrounding milieu.

The principle of homeokinesis assumes at its base, Claude Bernard's principle of the constancy of the internal environment as the conditions of free and independent life, (1859). We can take heart from the words of L. J. Henderson (1926), in commenting on this principle, that "This should not be thought of as absolute constancy, and it should be understood that variations in the properties of the internal environment may be both cyclical and adaptive, that is functional, but in general may not be random and functionless. Claude Bernard's principle is the first approximation which sufficies until the subject has been broadly developed". In our opinion, homeokinesis is the second approximation.

However, homeokinesis also expands the concept of dynamic mediation of periodic chains to include external environmental boundary conditions. Some chains are only closed internally, others externally.

A viable biological organism, in the words of Sechenov, ".... is inconceivable without an external environment for its existence ... a scientific definition of the organism should include also the environment by which it is influenced" (1861).

Spatial and Temporal Biospectroscopy - This is a matrix whose columns represent the chemical materials, the chemical process chains, and the complex behavioral modalities of the living system, and whose rows represent ordered time scales. In any column, the many temporal levels of regulation of that material, process, or modality are categorized.

It is our surmise that the mechanisms that exist behind this matrix represent the physiological - behavioral realization of Mackay's "conditional - probability matrix".[1] Conceived of in an information theory context, the

[1]See C. Cherry, "Information Theory", 1956.

conditional probability matrix may be viewed as the states and the frequencies of occurrence of these states to be found in a viable organism.

System Instability - The biosystem is marginally unstable. When at rest, it cannot remain at rest; when in motion, it cannot remain in motion. Thus, its behavior can only be stabilized as a steady state pattern of periodic variation among all of the major variables and states.

Behavioral Modalities - The result of this instability and the chains of organization that emerge, with both genetic and epigenetic links, is the development of a pattern of behavioral modalities, perhaps 10-20 in all. These modalities orchestrate the actions of the oscillatory chains around major melodic themes, the emergent 'hungers' of the organism. In the simpler organisms, these may be simple patterns of grow - divide, ingest, photosynthesize, move, eliminate, etc. In more complex organisms they begin to involve the externally linked psychological complexes of mating, child care, and play.

Hierarchical Regulation - As a corollary of the emergent behavioral modalities, the matrix of function that emerges within the complex biosystem is not simply spatially (morphologically) organized, nor temporally organized. Instead there is a complex of space - time structures and functions that tend to be tied together to exhibit the hierarchical categories or 'levels' of organization.[1]

The concept is still vague, but loosely, the following levels may be noted.

a. The individual cell, regardless of type, has its own life scheme within its operative environment, (eat from surround - grow - divide).

b. As a result of intercellular communication, specialized cells are yoked together to form a functional unit. This is reasonably well defined in organs.

c. The functional units are tied together to form an organ. It is the organ and its connections that create this organismic operational level. The whole organ probably very rarely, if ever, participates in its characteristic function, usually 1/2 to 1/4 or less of the functional units of the organ are involved.

d. The complex of organismic structures and their connections which are sufficient to make a biochemically operable system represents the lowest organism level. However, this has the characteristics of a lightly anesthetized animal, or animal in coma. It cannot persist without support functions.

[1]At any hierarchical level, there may be a 'heterarchy' or dissimilar structures or functions that are yoked together to provide the scenario or melody for that one level. We use the simile of orchestra and music because, in the melodic line, the heterarchy of instruments, the orchestra, and the orchestration, the complex of biological function is most easily discerned.

e. If we add to this system enough autonomic system function to cope with its external environment for the satisfaction of its major behavioral modalities, then we can identify this as a 'sentient' organism. This represents the reaches of most plant and animal species, or of a simple minded human being.

f. At the highest known level, we finally have the 'sapient' level, which is capable of interacting with and modifying its environment by the use of complex abstract internal languages. The essential sapient characteristic is likely the ability to translate among its many internal languages.

The study is then concerned with attempting to deal with some of the more rudimentary hierarchical levels in the biosystem within the framework of these principles, in particular the metabolic level.

After a long introduction (Part A) to the aforementioned physical ideas, in which physical scientist and biologist attempt to reach a mutual acceptance on how their points of view may be descriptively related to the biological organism, the biologist in turn (Part B) attempts to reexamine his biological ideas so as to conform to some extent with the physical description.

The first essay discusses the characteristics of the membrane as it is known and described within the current literature. It touches, in passing, on some of the current controversies that exist in this field. It would be difficult at present to develop a systems science that includes this level of biological action within an operational format. Yet the dynamics of the membrane sits on top of all biochemical, metabolic, and communications processes.

Electron microscopy has brought many of the apparent features of the membrane into view. It has not been able, as yet, to clarify and guarantee the transport and exchange processes. At present there is vigorous debate going on with regard to both passive and active transport through membrane walls of both small and large molecules. There is also an encouraging beginning at elucidating intercellular communications at the common membrane junction, both passive or active. It would only be speculative and risky for our group to attempt any detailed contributions to this biosystems level, other than an occasional physical analysis and commentary.

As an example of such commentary, a theoretical note is added describing how the everpresent fluctuating processes at a lower hierarchical level can be 'rectified' to provide active sources for transport of the succeeding level. This is suggested as the formal basis for active transport at the membrane level, and in principle, in extension, at every succeeding hierarchical level. In one view this represents the statistical mechanical or 'atomistic' basis for homeokinesis.

The second biological essay discusses the system at the level of the microvasculature. As far as possible it attempts to discuss the dynamics of form and the dynamics of function as viewed from the microcirculation system. It starts from the premise that it will attempt to deal with both materials transport and communications in vessels under 1/2 mm. diameter.

It discusses the dynamics of 'vasomotion', the changing diameter of small vessels, and leaves uncertainties as to the action of the precapillary level. It is suggestive that an actual picture of normal resistive control in active organs and tissues is wanting.

A corollary, an operational description of the nervous control at the microvasculature, is also incomplete. Here both morphology and the dynamics of the control action is in its infancy.

One very positive, basic thought that arises is the need to define the functional unit of each organ system. The functional unit comprises all the cellular components which make the organ unique. It includes the cellular elements of the organ, and the cellular elements that make up the microvascular supply and the nervous supply. The complexity of the defining task is illustrated for one functional unit, where it is moderately well known, the liver.

However, this essentially morphological description brings up the issue again of the dynamics of regulation and control at the capillary level, this time at a metabolic level. Here the issue tends to be chemical control of the very small vessels. Again a satisfactory agreement on both a description and theory of action is wanting.

Another unsolved problem is the moment to moment control of blood volume. It is presumed to be under the control of the venous system and the liver.

Finally, the essay touches on the difficulties of obtaining agreement on the actual structural picture of control of vascular permeability, the transport problem in the microvasculature. In one view, it is a static near-passive picture based on fixed pores; in another a more dynamic model is required. The essay ends on the note that in situ 'microscopy' in the living unanesthetized animal, preferably for a large animal such as the dog, should currently permit an adequate dynamic definition of some of the basic processes in a microvascular bed such as the functional unit of muscle.

The section on the microvasculature has two addenda, one that outlines the geometric - topological design in the microvasculature, and the second that illustrates some of the difficulties that lie in attempting to model the capillary exchange process. Water transport by the Starling model (outward filtration in the capillary entrance; reabsorption in the capillary exit by the interplay of hydrostatic and osmotic 'drives') is examined. There is difficulty in accounting for the Starling transport and for the presence of fixed pores.

The third essay presents an up-to-date view of how the autonomic nervous system may be related, by its innervation, to the peripheral organs. It uses the avenues opened by histofluorimetric study to trace out and speculate about the actions and sites of action of neurotransmitters. The autonomic ganglia begin to mediate the path for dynamic regulation by rate governing the input (i.e., it filters it). However, it is not clear how such filtered communications by neurotransmitters relates to the slower autonomic activity at end organs. Perhaps the end organs themselves mediate the release of

transmitter from the adrenergic terminals. Such possible 'feedback' control is discussed for the heart, the gut, a gland like the pineal, and the cardiovascular system more generally.

Besides the action of the cholinergic and adrenergic nerves (within the autonomic ganglia), the possible significance of chromaffin cells, lying in the vicinity of the ganglia in the many species studied, is outlined. They appear to act to influence ganglionic transmission, augmenting adrenergic function by mediating cholinergic transmission. The localization of chromaffin cells and their large catecholamine release in the vicinity of blood vessels provides a potential direct path into blood. How these cells may act as an externally distributed dynamic catecholamine release system, as they act in the adrenal medulla, is suggested for heart, corotid and aortic bodies, pancreas, thyroid, and in sympathetic prevertebral ganglia. The essay closes with the unsolved problem of accounting for a theory of the autonomic control of peripheral circulation.

The fourth essay discusses the cardiovascular system. First, it presents some of the primary properties of blood flow and metabolism, to establish the design relations that are likely common to all mammals. A weight specific nature of oxygen uptake and blood flow is suggested.

The developmental character of CV performance is outlined in hierarchical fashion. It is proposed then to develop an engineering description of the operational near-steady-state of the CV system. This is presented as the character of 400 second epochs of behavior. The rationale for this time state is presented as one for which near equilibrium metabolic events can be characterized; both the divisions of blood flow and the oxygen uptake will be near equilibrium. The description is developed as an engineering model of how the operating conditions of cardiac output and a 'grounded' venous pressure is achieved in the closed CV system. The necessity of relating the operating characteristics ultimately to microvascular parameters is again stressed.

The fifth essay is a most unusual attempt to bridge the gap from the biochemistry of the molecular biological level to the operative concepts governing the biochemistry at the organ level. It makes out a case for the dynamic homeokinetic model proposed in the introductory sections, and also for the need to focus on the functional unit in organs. It is clear that the organ element, the cardiovascular element, and the nervous element are all intimately involved in the biochemical chain.

The sixth essay is a rather extensive outline of neurophysiology. It was written to attempt to bridge the gap, as far as possible, between the communications characteristics of the nervous system and a synthesis of behavior - in the full periodic, episodic, and repetitious character that might be seen to emerge from the higher centers of the brain, notably from the cortex, reticular core, and limbic system. (That final synthesis we expected from Warren McCulloch. It is not written. We have, however, left the lengthy neurophysiological introduction.).

This essay starts from what is known about intercellular communications. It puts forth some of the modern views on how spatial - temporal patterns of

cellular differentiation develop. It discusses information transmission in the nervous system. It leans very heavily on sensory information, particularly visual.

It begins a discussion on how the nervous system achieves subgoals, 'sets' of the entire organism, as part of its hierarchical structure. The role of the reticular system is outlined, and of the autonomic system. The hypothalamic - pituitary link is discussed, and regulation through pituitary hormones is outlined. The combined action of nerves and hormones in controlling actions of the body is touched on. The essay ends with a discussion of the ontogeny of behavior. In particular it uses the genetic development of the visual system, from eye to brain, as a major illustration.

The last essay provides an introductory modelling of how behavior is related to the endocrine system. This final essay attempts to unify the description of behavior based upon neurophysical approaches with that based upon endocrinology and metabolism. It defines a modal view of human behavior and compares those modes with the smaller number of metabolic modes previously defined. It attempts to justify the postulate that all non-reactive human behavior is both modal and periodic, that the brain has a chemical basis for forming a body image of its structural and functional condition, and that the body image is partially dependent upon modal constellations of hormones and metabolites in blood. These constellations forming the body image are used by the reticular core and limbic system to urge decisions upon the cortex, and influence priorities among behavioral choices.

What has been clear to the contributers is that the task of putting the specialized subfields of biology[1] into a dynamic connected framework for form and function is inordinately difficult. However, for a first round of effort, the task is done. It will have to be redone a number of times, before it comes forth in a satisfactorily integrated form. However, this is not our immediate intention. Our problem is to try to decide, of the very many problems that remain, what problems can be fruitfully tackled.

We would like to furnish the following next keynote. The operation of a system can be pursued through its 'energetics', 'power engineering' or 'metabolic' processes, those that show the flow, distribution, and consumption of energy. In the main this has been our first round of effort. In a very rigidly yoked or coupled system, one that responds very mechanistically and deterministically to its environment - even if the system is autonomous - this would be the major method of tackling the system.

However, there is a second direction of pursuit, the 'communications' characteristics of the system. In this we include the complex of regulation and control by which a fixity of form and function manage to come off, even though the environment changes considerably. We believe that we may be up to the second kind of question.

However, for a first effort, we believe that too wide a scope to search out the information and communications characteristics of the complex biosystem would defeat us. Instead it is timely that we restrict our view. A particular system that may fit our talents and be of considerable immediate relevance is the reaches of the organism touched by the cardiovascular system.

Let us consider briefly what is required by a change (or upgrading) in point of view from the energetics to the command - control view of a system.

a. First, in order that the system be autonomously operative, it must be nonlinear and contain degradative sources.

b. Second, it must form autonomous thermodynamic engine cycles so that it can feed autonomously from the potential sources in the environment.

c. Third, if it is to persist autonomously in its environment it must have a command - control system that can store and respond to 'information' from the environment.

Thus, we are up to the question, what is information?

Information is a small amplitude energetic signal, patterned among all other signals in the environment, that has the capability of modifying the operative state of a system. The system must thus be capable of abstracting, storing and transforming these signals.

As we have preceded, hierarchically up the biosystem's levels, at the

[1]Confined in this report to membrane, microvascular, nervous autonomic, cardiovascular, hormonal, higher nervous structure and function.

highest 'sapient' level, the capability of storage and transformation is the greatest. In fact we have suggested that the 'sapient' character is to be associated with the ability to transform by 'translating' signal among many modes.

Thus, in developing an information theory view of the biosystem - in particular in the languages of the cardiovascular system - we shall have to begin to introduce the following language, which may appear quite strange to biology.

language - code, signals, signalling element, sign, symbol, character, letter, phoneme, word, sentence, message, sequences, dictionary, repertoire.

noise

transmission capacity

amplifier

receiver

sender

encoder

translator

transponder

transducer

channel, band width

information source

transmitter

message entropy

redundancy

Markov chain

pattern recognition

logon, metron

semiotics - syntactics, semantics, pragmatics

cybernetics

Roughly, as an introductory paragraph, we can view the problem of the cardiovascular system in the following light. There is an autonomous limit cycle oscillator (the S - A node, with back-up systems should that fail). This acts as a thermodynamic engine utilizing small power. This oscillator is used as an escapement on a larger oscillator, the intermittently excited heart. The heart, supplied by its power system, acts as a thermodynamic engine.

This heart engine pumps blood. It does so within the body. However, it can perform autonomously outside the body too. Within a section of this report, we have outlined its nominal mechanistic characteristics; how it responds to quasi-static changes in motor activity level of the organism; and

how it changes adaptively to longer term changes in motor activity. We have thus discussed both its source and receiver characteristics, as an energetic system.

We have not discussed, however, its information theory characteristics. This is how it responds to 'communications' signals, both transient and behavioral; how it responds to the changing visissitudes of the milieu. We do not know what languages it is communicated with, nor to which it transponds. We have to describe the 'language' of the exchange processes.

All of these tasks are thus designed to try to bring the active processes in the cardiovascular system, in the range of small signal chemical, catalytic, and chemo electric characteristics, within the scope of a communications theory description. What are the chemical and electrical languages of the heart and the systems connected to the heart?

XIII. ACKNOWLEDGEMENT

In addition to the support provided by the National Aeronautics and Space Administration under contract NASW-1815, the added support of the Army Research Office, funded under the same contract is gratefully acknowledged.