NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us a poor photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30. Please read the authorization forms which accompany this thesis.

THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de mauvaise qualité.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30. Veuillez prendre connaissance des formules d'autorisation qui accompagnent cette thèse.

LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS REÇUE

NECESSARY TELEOLOGY

by

Thomas Edward Fenn

Thesis submitted to the School of Graduate Studies and Research
of the University of Ottawa
as partial fulfillment of
the Master of Arts Degree (Philosophy)

NECESSARY TELEOLOGY TABLE OF CONTENTS

CHAPTER I

THE DEBATE ABOUT TELEOLOGY

1. INTRODUCTION

Theoretical inquiry into biology has developed around two primary interests: determining the logical organization of biological behaviour and therefore, as well, specifying the logical conditions for knowledge about biological organization. Each interest has produced separate but related issues in Philosophy of Biology. First is the issue about whether all organic activities are exemplifications of the same laws as those which govern non-living systems, or whether the purposive and goal-directed behaviour of organic systems indicates a special mode of causation in such systems. Second is the issue over the wide use of teleological explanation in modern Biology. Teleological explanations typically ascribe some function or purpose to an item or process and argue from the occurrence of the function to the presence of the item functionally characterized. The problem is, such explanations entail the questionable assumption that biological items and events are due to the purpose or design that is served by them. Yet if we reject teleological explanation as having matters in reverse, we are left wondering whether ordinary causal explanation is sufficient to accounts of biological organization, and thus returned to pondering a special mode of causation. Together these issues form the core of the debate about teleology among biological theorists. Of course, the positions taken have foundational significance for empirical research. They provide the theoretical framework for understanding what we disclose as we investigate living systems, and they shape our perception of what it is we are doing, or should be doing, in Biology.

My thesis is an argument for the use of teleology in the biological field. Biology's main explanatory science, physiology, leaves the field rich in teleological phraseology and explanation. My defence rests on an understanding of the work of Sir Charles Scott Sherrington (1857-1952), the eminent Briton whose exhaustive study of the spinal reflex culminated in his landmark theory of the integrative action of the nervous system. I try to develop his physiological theory in Chapter II. My argument is, first, that Sherrington's theory of integrative action, which champions the use of teleological concepts in physiology, is superior to the influential theory of self-regulation developed by philosopher Ernest Nagel, which unsuccessfully attempts to characterize purposive behaviour in strictly causal terminology. Chapter III contains this argument on the first issue in the debate about teleology. On the second issue, I concentrate less on the end-product of Sherrington's research, the integrative action theory, and draw more on the method of his research and explanation to show how we can make sense of teleological or *functional* explanation as it arises in, particularly, the biomedical field. My argument on the second issue is the subject of Chapter IV. I believe the argument resolves Carl Hempel's well-known trilemma whereby functional explanation is concluded to be either non-deductive, non-empirical, or else trivial. Yet I am equally convinced that functional explanation of the sort found in the biomedical field conforms to Hempel's *deductive-nomological* model of explanation, and I argue against the view of Robert Cummins that functional explanation incorporates a radically different logic from the one which Hempel supposes.

The remainder of the present chapter is an account of the positions of, first, Nagel and then of Hempel and Cummins. Chapter V is a brief conclusion that teleology is essential to physiology, and therefore to the biomedical field, but that its use does not breach our materialist presuppositions about empirical science. Lastly, I have included an appendix which discusses what is wrong with Sir John Eccles' view that Sherrington was a mind-body dualist; Eccles is one of many famous Sherrington students.

2. THE FIRST ISSUE: UNDERSTANDING PURPOSIVE BEHAVIOUR

2.1 PRELIMINARY STATEMENT

Everyday experience of plants and animals leaves one impressed with nature's orderly process. Sweet flowers open and bees collect their food. Excretion normally occurs a predictable number of hours after the consumption of food. Societies get organized as labour becomes divided. Spermatazoa are motile. Natural processes such as these are not haphazard, and by reason of the complexity of their relationships and design, they even appear to be more than merely causal. A characteristic feature is the organization of such processes around nameable goals or purposes. In the above examples, reproduction, consumption, digestion, and syngamy are the respective purposes served by the organizational features of each of these natural processes. Biologists have uncovered further purposive mechanisms which *control* metabolism, temperature, blood-circulation, hormonal change, and reflex action, for example.

Purposive organization is found to be the most characteristic feature of biological behaviour at all levels, from physiology to society and natural economy. The problem is to understand how matter must be organized to accommodate purposive behaviour.

2.2 TWO EARLY VIEWS

The occasionalist view of Nicholas Malebranche (1638-1715) blended theology with speculation to raise the purposiveness observed in the development and behaviour of plants and animals to expressions of God's will at a distance. The phenomenal symphony is created by the Supreme Being, and depends upon Him for its existence from moment to moment. God constantly orchestrates the actions of the billions of mindless biological processes and individual organisms. Biological organization and natural economy are purposive because God created the entire universe, and He does nothing without a purpose.

Malebranche's view posed many problems for the development of rational answers to a variety of questions, not the least of which, for example, is the question of man's free will. How is free will possible, it would be asked, if God creates and controls everything in the universe? In science, moreover, God's omni-competence entails that every event is an expression of divine reason or purpose, which makes it impossible to draw the distinction observers could reasonably draw between the purposive behaviour of biological events, and the ordinary causality of physical events. Water expands when it freezes, it would appear, because of its chemical properties, not because an ulterior plan is being served. By contrast, a peacock will strut around inflating its

throat and fanning its feathers, clearly to attract the attention of a nearby peahen. Compare the body temperature of a corpse with that of its pathologist, and one finds the former varies directly with the ambient temperature, whereas the pathologist enjoys an automatically adjusting and fairly constant internal temperature, quite independent of the same environment. Or examine the hide of any dog about to exercise its scratch-reflex and one very often uncovers a flea, the removal of which is the purpose of that bit of spinal behaviour. Malebranche's view failed to support the distinction that would eventually appear to separate biology from physics, but worse yet, it undermined the possibility of scientific knowledge altogether. If the explanation for all phenomena, organic and physical alike, rests with an understanding of God's will, and if knowledge of God is ultimately impossible, then of course no knowledge of the world is possible.

Through the interest and effort of men such as Rene Descartes (1598-1650), empirical inquiry persisted. The idea arose that what can be observed may be explained by understanding something of the mechanics of the things themselves. Thus, in the biological field, attention shifted from unprofitable speculation on a Final Cause, to a quasi-empirical search for a *special* cause. Descartes supposed that each individual organism is the residence of a "dynamic agent", a non-physical entity having causal efficacy which produces purposive behaviour. The ethereal "animal spirits" surge through any of a number of observable vessels or channels which ramify throughout the body, and form a type of hydraulic

transmission system. The animal spirits must have an innate 'awareness" of the purpose of every bodily process and behaviour, and and innate ability to operate the body mechanism to bring about any proper result or state of affairs. The spirits could influence the body with or without a cohabiting soul or mind, although it was obedient to the will if the organism had one to exercise. The purposiveness Descartes observed in the coordinated spinal reflex was explained in terms of the hydraulic activity of animal spirits. The causality of purposive systems so conceived can be considered special in so far as the effects produced appear to be influenced by goals which lie in the future and goals of which the animal spirit is "aware". Ordinary causality, on the other hand, works from past or present conditions only. Physical events are "pushed" into existence by the past or present, if you will, whereas teleological ones so conceived, are "pulled" into the present by an "awareness" of the future. The outcome of the scratch-reflex, for example, was in large measure seen to affect how the reflex operated.

As biological science grew more clear about itself, however, teleology became less credible. Proof and testability became conditions of scientific truth generally, and no less so in biology. The trouble with teleology, viewed as special causality, is that once again experimental testing is impossible. That which has yet to exist -- some future goal -- cannot be demonstrated to affect any present set of observable conditions. In contrast, ordinary causal explanation develops from observing the outcomes of conditions which can be experimentally recreated for demonstration

and proof. The idea of teleology as special causality probably arose from a conflation of biological acts with human acts, and then from a misunderstanding of the causality of intentional actions. Biological acts are like human acts because they are purposive, and human acts are accompanied by mental processes, so why not suppose that biological acts are accompanied by, at least, mental-like processes as well? And again, goal-directed mental behaviour might be mistaken for proof that the future can affect the present. Goals are things towards which we strive, and they lie in the future; therefore the future affects the present. But of course, this argument is based upon a faulty analysis of the causality of human action, and ~~is thus~~ unsound. What is causally efficacious is not a future, yet-to-be-realized goal, but "(a) (the) desire, present before the action, to attain that particular objective, and (b) the belief, likewise present before the action, that such and such a course of action is most likely to have the desired effect." (C.G. Hempel, *Aspects of Scientific Explanation*, 1965) Thus teleology as special causality is empirically vacuous either because it refers to entities which cannot, in principle, be observed, or else it makes reference to conditions which do not admit of empirical testing. The result, in either case, leaves biological events ultimately imponderable once again.

2.3 MODERN SCIENCE AND PHILOSOPHY ON PURPOSIVE BEHAVIOUR

This is largely where the problem of teleology sat when Sherrington came to the scene.

> Older writings on reflex action concerned themselves boldly with the purposes of the

> reflexes they described. The language in which they are couched shows that for them the interest of the phenomena centered in their being regarded as manifestations of an informing spirit resident in the organism, lowly or mutilated though that might be. Progress of knowledge has tended to unseat this anthropomorphic image of the observer himself which he projected into the object of his observations. The teleological speculations accompanying such observations have become proportionately discredited. Self-wounded in this way physiology became for a time extremely reticent about purpose, remaining simply objectively descriptive.

Sherrington's lifetime ambition was to *explain* how purposive, coordinated behaviour in the spinal reflex operates. Before him lay only scattered attempts to understand the physiological mechanism of the nervous system. The complete picture was unrecognizable. He relied mainly upon his own exhaustive research to piece together the puzzle of reflex action, and he knew that he must find a way to use teleological concepts, without falling prey to the pathetic fallacy he referred to in the above passage, if his efforts to *explain* reflex action were to prove successful; without teleology, in Sherrington's view, physiology was merely a descriptive science. As the passage continues:

> The impetus given to biology by the doctrine of adaptation under natural selection, felt so strongly by morphological studies, seems hardly as yet to have begun its course as a motive force in physiology. But the signs begin to be numerous that such an era is at hand. The infinite fertility of the organism as a field for adapted reactions has become more apparent. The purpose of a reflex seems as legitimate and urgent an object of natural inquiry as the purpose of the colouring of an insect or a blossom. And the importance to physiology is, that the reflex action cannot be really intelligible to the physiologist until he knows its aim.
> (C.S. Sherrington, *Integrative Action of the Nervous System*, 1906, p. 237)

Sherrington's theory of the integrative action of the nervous system provided the conceptual framework for subsequent studies of the brain, as well as the basic model of purposive mechanics for the rest of physiology. With physiology as the main explanatory science in the field of biology, the importance of Sherrington's theory to the history of ideas cannot be over-stressed. Equally apparent is its importance to the philosophic debate on teleology.

When philosophers have examined Biology, the usual ulterior motive concerns the question of autonomy. Knowing the logical organization of biological behaviour entails the problem of the relation between such organic activities as growth, regulation and reproduction and the complex physical and chemical processes that go on in living systems. (*The Encyclopedia of Philosophy*, "Biology", 1967) To the mechanist, all organic activities are governed by the same laws as those which govern non-living systems. If this is true, biology should become ultimately a branch of the physical sciences, and have no methodological autonomy. The methodological unity of science is the thesis Nagel wants to protect in developing his view that teleological concepts can be analyzed into ones of efficient causality.

In contrast with Sherrington, Nagel holds that teleological concepts are to be avoided, not regarded as epistemologically necessary to physiological understanding and explanation. After analyzing his model of self-regulation, Nagel concludes:

> The ... analysis has therefore shown that the notion of a teleological system can be explicated in a manner not requiring the adoption of teleology as a fundamental or unanalyzable category.
> (E. Nagel, *The Structure of Science*, 1961, p. 417)

If Nagel is right, the use of teleological concepts in biology is no bar against the reduction of biology into physics.

If, however, teleological concepts cannot be so analyzed, we must question seriously their meaning and epistemological value. Moreover, teleology is no longer generally regarded as a special causality, but a *special organization* of ordinary causal processes. This sits well with both Sherrington and Nagel, but I believe Nagel's ulterior motive, to protect the methodological unity thesis, has caused him to ignore or else overlook a significant development in physiological thinking since Sherrington gave that science a new direction. That is, Nagel fails to explore the possibility of understanding purposiveness in and of itself, as it were, without reference to underlying causal processes. As a matter of fact, however, this possibility is precisely what modern physiologists are exploiting, in their effort to quantify the effects of control mechanisms; mechanisms, in other words, which form the behaviour we have described as purposive. I will examine this development more closely in my third chapter. I should merely point out presently that, I believe the effort of these physiologists can continue confident in the knowledge that teleology *is* a fundamental and unanalyzable category necessary to understanding and explanation in physiology.

2.4 NAGEL'S THEORY OF SELF-REGULATION

The teleological system Nagel cites as an example of his theory of self-regulation is temperature homeostasis, or the maintenance of constant internal temperature, in the human body.

As with electrolyte and water homeostasis, blood-sugar homeostasis, blood-oxygen homeostasis, and so on, temperature homeostasis is a relatively steady state maintained "by complicated but coordinated physiological processes." (409) The avoidance of fatal injury requires the maintenance of a fairly constant internal temperature in the presence of sometimes extreme variations in the temperature of the environment. During the normal day, body temperature varies "from about 97.3°F. to 99.1°F",Nagel reports, and otherwise must be maintained within the restricted range of values between 75°F. and 110°F. to avoid death, and to enable the organism to carry on with whatever activity it happens to intend. In being able to maintain temperature homeostasis, the body can continue with its normal activities "in relative independence of the environment." As Nagel explains:

> The body achieves this homeostasis by means of a number of mechanisms, which serve as a series of defenses against shifts in the internal temperature. Thus, the thyroid gland is one of several that control the body's basal metabolic rate (which is the measure of heat produced by combustion in various cells and organs); the heat radiated or conducted through the skin depends on the quantity of blood flowing through peripheral vessels, a quantity which is regulated by dilation or contraction of these vessels; sweating and the respiration rate determine the quantity of moisture that is evaporated, and so affect the internal temperature; adrenaline in the blood also stimulates internal combustion, and its secretion is affected by changes in the external temperature; and automatic muscular contractions involved in shivering are an additional source of internal heat. There are thus physiological mechanisms in the body that automatically preserve its internal temperature, despite disturbing conditions in the body's internal and external environment. (409)

One could add to my list of self-regulating systems which maintain a homeostatic state, the scratch-reflex characterized by Sherrington. A disturbing irritant which stimulates certain receptors in the hide of the host animal will cause a chain of neurones, leading from the receptive surface to the hind limbs, to become excited. The neuronal activity causes the limb to scratch the spot and remove the irritant, thus returning the nervous system to its basal or homeostatic state of excitation. For Nagel, the important point is that self-regulating systems have the capacity to compensate for a restricted range of distrubances to maintain, or return the system to, a homeostatic state.

The distinguishing features of a self-regulating system become part of the topic of Nagel's later essay entitled, "Teleology Revisited" (*Journal of Philosophy*, May 1977). A self-regulating system has three main properties, according to Nagel; it is:

a) plastic
b) persistent
c) orthogonal

a) These systems are said to be *plastic* "in the sense that the goal can be reached by alternative routes and from different initial starting points, the route or action that is actually adopted depending upon what local circumstances prevail." (265) In the example of temperature homeostasis, the options available to the system to achieve the goal of keeping a constant internal temperature are considered to be: activation of the thyroid gland; control of heat loss through the skin by peripheral vasodilation and vasoconstriction; exothermic moisture evaporation through

sweating or panting; discharging adrenaline into the bloodstream to increase the metabolic rate and thus the production of heat; activating the shivering response to cause a great increase in the quantity of heat produced. Presumably, the nature and extent of any internal or external disturbance determines which one, or set, of these routes becomes operative in maintaining temperature homeostasis in a given situation. Of course, one would seek the advice of a physiologist, not a philosopher, on matters such as this.

Yet there is an important philosophic point which should not go unnoticed. If the maintenance of temperature homeostasis, or the homeostasis of any set of conditions, is a function of the activities of all mechanisms working in unison, what sense can be made of the claim that self-regulating systems are plastic? Maybe, as a physiological fact, the components of a physiological system are so interdependent, or mutually influential, that no single pathway, or mechanism, can be so isolated and named as anything more than *contributory* of the overall effect on the system as a whole. In other words, plasticity may have methodological significance, but no real basis in physiological fact. Methodologically, it may be necessary to isolate pathways and mechanisms in order to assist psychological understanding of these complex systems. Plasticity would then be considered to have heuristic value, but no epistemological significance with regard to the real biological world. Yet if self-regulating systems are not in fact plastic, the heuristic value of the concept would be highly questionable. It could lead us from and not toward an understanding of such

systems. Analyzing a system into its parts is a step toward being able to describe the system as nothing more than a collection of parts -- a step toward reductionism to protect the methodological unity thesis; but, at least in the case of the self-regulating central nervous system, such a step is not in the direction of a final and complete understanding the theory of intergative action, and how the nervous system works to produce purposive, coordinated reflex action. Anatomical plasticity is not a property of the central nervous system, a self-regulating system *par excellence* as we will see, and plasticity would, in fact, unsurp the possibility of purposive, coordinated reflex action. Moreover, an analysis of the whole into its parts, which the concept of plasticity requires, is not the final step toward understanding that system. An understanding of the integrative action of the nervous system does not end with conceptually disintegrating that system, as it were. And to repeat the point I made earlier, modern physiologists are concerned to understand the effects of control systems *qua* system, as well as, of course, the special effects of any component on the whole. So the application of the concept of plasticity may be seriously limited in physiology. It is at least true that, logical analyzability must not be mistaken for physiologically isolated pathways and mechanisms.

b) The next most characteristic feature of self-regulating systems which Nagel describes is *persistence*. By "persistence" he means simply that if disturbed, the system has the capacity to compensate

and reduce the effects of the disturbance, thereby returning the system to its homeostasis. The five mechanisms or pathways to thermal homeostasis Nagel described enable the body to be persistent with respect to the homeostasis of its core temperature. Changes in either internal or external temperature, provided they are not too great, provoke compensatory changes within the system, which return it to or maintain a temperature around the optimum value. The important point is that homeostasis is possible as a result of certain persistent changes occurring within the system itself. Without persistent compensatory changes, the system would fail to maintain homeostasis in situations of internal and external disturbance.

c) Systems which are self-regulating with respect to some goal (e.g. thermal homeostasis) are distinguished from those which are not, according to Nagel, by the characteristic that the variables of the former are *orthogonal*, or independent of one another. Nagel considers orthogonality to be the formal or logical criterion which distinguishes the two groups of systems. In the case of temperature homeostasis, two variables, say the blood-thyroxine level and the blood-adrenaline level, are independent in Nagel's sense because the values assignable to the one variable are not *logically* determinant with respect to the values assignable to the other variable. That is, the value of one is compatible with all values of the other. It is important that the variables be independent in this sense, Nagel believes, otherwise the values of one could be deduced from the values of the other. If this held, Nagel implies, the system could be characterized in terms

of the determinant variable alone, since it would be redundant to mention the second variable when defining the primary properties of the system at hand. The criterion of orthogonality supposedly safeguards against equating physical processes of simple equilibrium with systems of self-regulation. Consider for example the physical system of a beaker of water of 10°C. placed into a container with an ambient temperature of 0°C. Heat loss from the beaker of water will raise the ambient temperature, while reducing the water temperature. The two variables in the example are the temperature of the water and that of the environment. Clearly, the former value is inversely related to the latter value, and vice versa, for all states of equilibrium. That is, raise or lower the initial ambient and water temperatures and, *ceteris paribus*, the resultant thermal equilibrium value is the product of the cancelling effects of each variable on the other. Similarily, as Nagel states, the criterion of orthogonality precludes the flow of a river from being viewed as a goal-directed, or self-regulating, process. Although the river flow persists towards the sea and is persistent in the face of a wide range of obstacles, the flow is not self-regulating because its direction depends upon the lie of the land. Basing the distinction bewteen these two kinds of systems upon Nagel's concept of orthogonality is highly questionable as we will see. Plasticity, persistence, and especially orthogonality are, none the less, the three properties which uniquely characterize self-regulating systems for Nagel. These properties which characterize purposiveness, or goal-directedness, are explicable in terms of the dispositions of the components of such a system. If the analysis is sound, teleological behaviour is explicable in terms which entail

no teleological concepts, and the reduction of biology into physics becomes possible.

We now return to the earlier essay in Nagel's *Structure*, and to the problem of explaining purposive behaviour in terms which are strictly causal, and do not imply ones which are teleological. Using the preceding discussion as the intuitive context of Nagel's explanation, I will attempt to develop his concept of goal-directed behaviour axiomatically.

Nagel equates teleological systems with goal-directed, directively organized or self-regulating systems and renders what he believes to be a purely causal analysis of such systems. As indicated above, the organization of these provides that, if disturbed by some internal or external cause, the set of altered conditions are so causally efficacious as to further alter certain other conditions so as to return the system to its original state, a state referred to as "homeostasis". Nagel would explain that owing to the causal efficacy of the system upon itself, such systems appear to be directively organized toward maintaining homeostasis. That is, they may appear to operate by means of teleology as special causality where future events affect present conditions, but, in fact, Nagel would say, all of the properties of purposive behaviour are explicable in terms of the present and past dispositions of the components of the system. From this, he reasons that there are no uniquely biological properties of purposive systems, and thus no uniquely biological laws must be invented to explain their behaviour. The methodological unity of science is therefore uncompromised in Biology. Between pages

410 and 421, Nagel formalizes his concept of self-regulation. Its logical development may be represented as follows.

I. Definitions

1. Let S be a directively organized system
2. Let E represent the environment
3. G is a state, property, or mode of behaviour which S is capable of possessing, or maintaining (e.g. a body temperature of 98.6°F.)
4. A, B, and C denote parts of system S (for simplicity, assume there to be three parts or mechanisms; e.g., for temperature regulation in the human body, assume the available pathways are: thyroid control of metabolism; peripheral vasodilation and constriction; and exothermic moisture evaporation through sweating or panting.)

II. Properties of System S (necessary conditions)

1. Each part is indispensible
2. Each part is capable of more than one state
3. The set of possible states of each part is a restricted class (K_A, K_B and K_C)
4. The value (state) of any part is logically independent of any other at the same moment (*orthogonality*)
5. The parts are so interrelated that any change in one produces compensatory changes in the remainder to offset the effects of the disturbance (*persistence*)
6. The disturbances to which the parts are susceptible are such that compensatory changes can be initiated by any of the parts of the system (*plasticity*)
7. The whole (system S) is deterministic with respect to its parts (i.e., these and only these parts determine all possible G-states)

III. Parameters of behaviour of System S

1. In principle, S may have more than one G-state
2. In principle, G may be unattainable

3. If G is attainable, then either S has property G, or S will have property G
4. Not every possible state of S is a G-state
5. If S has more than one attainable G-state, then the actual G-state at time t_1 is determined by the actual S-state of some previous time, t_0

The five parameters of behaviour, or possible types of behaviour, are determined by the properties of system S. Parameter III, 1, that S may have more than one G-state, follows from properties II, 2, 3 and 7, plurality and class restriction of part values, and parts-whole determinism, respectively. Given that the overall disposition of system S is directly related to the properties (values) of the parts (II, 7), variations (II, 2) in the latter will permit variations in the former. Equally, restrictions (II,3) upon variations in the latter will mean restrictions for the former. Thus, we can expect to find a range of possible G-states for system S. In the example of thermal homeostasis, the range of normal core temperatures falls between 97.3° F and 99.1°F and, in theory therefore, the possible G-states are infinitely many.

S can be seen to have more than one G-state for the additional reason that there exists a *set* of possible states for the three parts, or varying degrees of involvement of each mechanism, and any possible state collectively characterizes a possible G-state. Since the description for any possible state will, by definition, differ from that for any other, the collective characterization will differ as well. The idea is similar to considering the possible combinations of any three variables, the numerical sum of which equals any number you wish, say 10, where 10 is analogous to a

particular G-state. The cominations, 1, 2, 7; 2, 7, 1; and 7, 1, 2 all equal 10, although by virtue of different series of numerical values (parts). Similarily, any G-state for Nagel may be the outcome of different combinations of the effects on system S of the, theoretically, three parts.

Parameter III, 2, that certain G-states will be impossible to attain, also follows from properties II, 2, 3, and 7. Given the restrictions placed upon the parts, certain goals, or G-states, will be unattainable.

Parameter III, 3, that the possibility of attaining G implies that S either has G or will have G, is a function of the interrelationship between the parts (II, 5), about which more can be said shortly.

Parameters 4 and 5 follow from properties II, 5 and 6. Straightforwardly, if it is possible for S to attain G, and if S is not in a G-state at time t_1 then, at t_1, S is characterized by a state which is not a G-state. Secondly, if all possible G-states are determined by the possible values of, or relations between, the parts, any actual G-state is determined by any actual relations between the parts. Since the relations are always such as to return the system to a G-state at time t_1, assuming that system S is not in a G-state at time t_0, whenever a particular G-state is maintained, it is necessarily the result of some previous S-state.

How these properties of system S emerge, the most characteristic of which are presistence, plasticity and orthogonality, is explained in terms of the reciprocal relationships the parts have

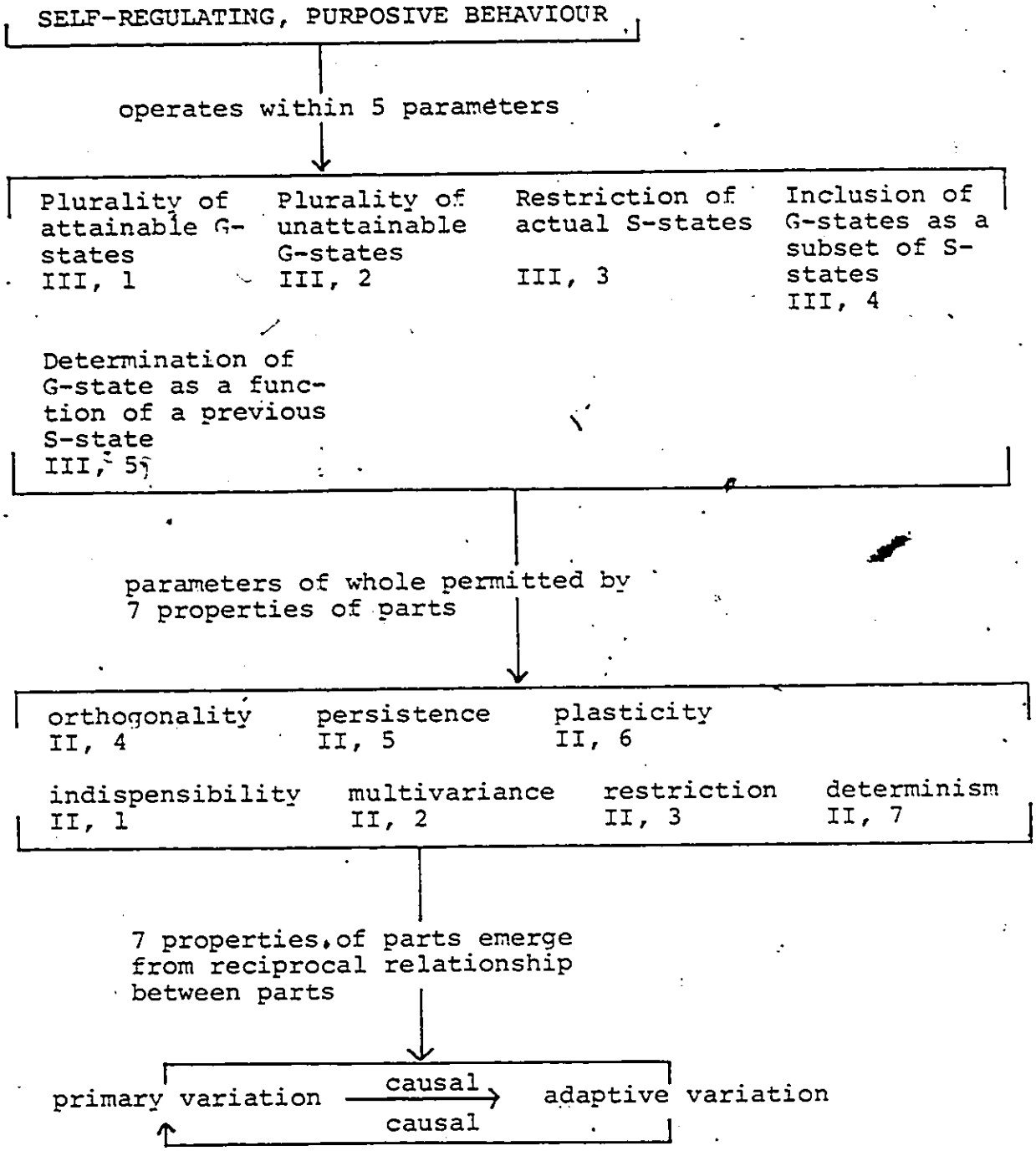
to one another. At this point in his discussion, Nagel introduces the concept of a *primary variation* which evokes *adaptive* or *compensatory variations* in the remaining parts, which adjust the system and return it to the G-state. He writes:

> If S is in a G-state at a given initial instant t_1 falling into some interval of time T, a change in any of the state variables will in general take S out of the G-state. Assume that a change is initiated in one of the state variables (say the parameter 'A'); and suppose that in fact the possible values of the parameter at time t_1 within the interval T but later than t_1 fall into a certain class K_A, with the proviso that if this were the sole change in the state of S the system would be taken out of its G-state. Let us call this initiating change a "primary variation" in S. However, the parts A, B, and C of S are so related that, when the primary variation in S occurs, the remaining parameters also vary, and in point of fact their values at time t_1 fall into certain classes K_B' and K_C', respectively. These changes induced in B and C thus yield unique pairs of values for their parameters at time t_1, the pairs being elements of a class K_{BC}'. Were these changes the only ones in the initial G-state of S, unaccompanied by the indicated primary variation in S, the system would not be in a G-state at time t_1. As a matter of fact, however, the elements of K_A' and K_{BC}' correspond to each other in a uniquely reciprocal manner, such that, when S is in a state specified by these corresponding values of the state variables, the system is in a G-state at time t_1. Let us call the changes in the state of S induced by the primary variation and represented by the pairs of values in K_{BC}' the "adaptive variations" of S with respect to the primary variation of S (i.e., with respect to possible values of the parameter 'A' in K_A'). Finally, when the system S satisfies all of these assumptions for every pair of initial and subsequent instants in the interval T, the parts of S causally relevant to G will be said to be "directively organized during the interval of time T with respect to G" -- or, more shortly, to be "directively organized", if the reference to G and T can be taken for granted. (415)

Adaptive variations thus counteract the effects of primary variations on the system, such as to adjust the system in a manner which maintains a certain internal state of affairs (G-state), giving the system the appearance, over time, of being directively organized, self-regulating or purposive. The reciprocal relation between the parts requires that the parts be: indispensable, otherwise the causal countereffect necessary to self-regulation could not obtain; multivariant, to permit a range of possible adaptive variations; class-restrictive, to enable any adaptive variation to be truly adaptive and not counterproductive; orthogonal, to allow the physiological interaction which occurs between the variables; persistent, in so far as the overall effect on the system of the counterplay between primary and adaptive variations is to maintain a G-state; plastic, to enable the system to respond appropriately to a wide range of possible disturbances; deterministic, to allow the system to be exhaustively characterized in terms of the possible values of its variables. The following chart should help clarify the analysis, which develops from the phenomenon of self-regulation to the causo-physical interplay between parts, variations in the values of which are either primary or adaptive. (See chart page 23).

Ultimately, I believe, Nagel's distinction between systems of self-regulation and ones of simple physical equilibrium collapses, unless we make use of teleological concepts. In other words the concept of orthogonality, the most characteristic feature of self-regulating systems for Nagel, logically requires teleological concepts for meaning. Moreover, the reciprocal relationship between parts

Chart I



The analysis proceeds from self-regulation to the reciprocal relationship between parts; properties II, 4, 5, and 6 are major, the remainder are presupposed.

which gives rise to this most characteristic feature, and which supposedly provides a purely causal account of goal-directed behaviour, is in fact covertly teleological. It too logically requires teleology for its explication. If the distinction between self-regulating and equilibrium systems is useful and meaningful in science, as I believe it to be, teleological concepts are also useful and meaningful. And if the idea of a reciprocal relationship between components in a system is physiologically significant, which it is, then so too are teleological concepts. Just how teleological concepts enter the picutre Nagel has given us will form part of my critical appraisal in Chapter III. We can now proceed with the second issue in the debate on teleology.

3. THE SECOND ISSUE: THE LOGIC OF FUNCTIONAL EXPLANATION [1]

3.1 SCIENTIFIC EXPLANATION

---

[1] Bibliographical note: This section concerns the work of two philosophers, quotations from whom are followed by a page number enclosed in brackets. The works are: ASPECTS OF SCIENTIFIC EXPLANATION, C.G. Hempel, The Free Press, New York, 1965; and, "Functional Analysis", Robert Cummins, THE JOURNAL OF PHILOSOPHY, 20, November 1975.

Scientific explanation is a subject central to philosophy of science and no less so in the special field of philosophy of biology. It is held that research in science aims beyond the discovery and description of natural events and processes. Understanding is the target, understanding *why* phenomena happen as they do, when they do. Explanation is taken to be the criterion of scientific understanding. Hence specifying the logic of scientific explanation is a philosophical problem. Yet how the problem is conceived depends on whether it is taken to be a criterion of understanding in the sense of a successful resolution of someone's psychological puzzlement about a matter, or a criterion for a formal concept, the satisfaction of which entails conditions such as conceptual completeness, parsimony, consistency and proof. Moreover, offered solutions to the problem of scientific explanation articulate different views on cognitive significance, the nature of a theory, the place and duty of a law, and even about the verity of things perceived, and whether the true aim of science is not just understanding, but control. The problem is complex, and many issues can be covered. For this reason, philosophy of science is considered a fresh perspective on many traditional problems of philosophy.

My interest in this section is restricted to two opposed positions which do consider explanation to be a criterion of formal understanding. These positions profess to be precise accounts of the logical structure of explanation. Accounts of this sort enable us to show whether the primary aim of a given science, as formal understanding, is ever achieved; which of its purported explanations

fail, and why, and where formal understanding is logically weak. The science in question is of course biology.

3.2 PRELIMINARY STATEMENT

Explanations in biology characteristically use the term "function" in accounts of physiological and behavioural processes. For example:

(1) The function of the heartbeat in vertebrates is blood circulation.

(2) The function of the contractile vacuole in fresh water protozoa is the elimination of excess fluid from the organism.

(3) The kidney's function to remove electrolytes, water and waste products from the blood stream.

(4) The scratch-reflex functions to groom the hide.

Frequently, the notion of function is expressed by equivalent terms, such as "serve" and "role", and sometimes it is simply implied in a statement. Thus:

(5) Piloerection serves to instil fear.

(6) The colourful plumage of male birds has a key role in mating.

(7) The anthropoid brain guarded against the extinction of otherwise defenceless primates.

What the notion means is problematic, as is whether its use affects the logic of functional explanation. By contrast, physics and chemistry explain phenomena in strictly causal phraseology that does not, implicitly or otherwise, rely upon the notions of function purpose or goal. This has lead to the suspicion that biological science is methodologically autonomous.

Such a position has been developed by Robert Cummins in his essay "Functional Analysis". Cummins has many strong objections to Carl Hempel's presumption that all explanation in science should conform to his deductive-nomological model (hereafter referred to as the D-N model) and that functional explanation in biology is no exception. I will examine the D-N model, and then how Hempel extends its application to functional explanation. Cummins' criticisms are to be discussed and his alternative model of functional explanation will be considered. I hope to offer a reasonable defence of Hempel's core position on the D-N model and show why Cummins' positive thesis fails. Chapter IV contains what I think is a more satisfactory account of the deductive-nomological nature of functional explanation.

3.3 THE DEDUCTIVE-NOMOLOGICAL MODEL OF EXPLANATION

"The Logic of Functional Analysis" begins by restating the well known model of D-N explanation. Hempel designed this model to account for properties necessary to adequate explanation in an exact science. Physics and Chemistry are his exemplars. Any purported explanation meets the requirements of the D-N model, and qualifies as scientific explanation, should it:

(1) Make explicit reference to one or more general law(s),

(2) Involve a deduction of the *explanandum* from the *explanans*,

(3) Have empirical content, significance or meaning. (247)

The idea is thus represented:

explanans	1. $L_1 \ldots . L_n$	- statement of one or more general law(s).
	2. $C_1 \ldots . C_n$	- statement of initial and boundary conditions
explanadum	3. Therefore E.	- statement of the event to be explained (usually of the form, "Therefore E occurred.")

A general law specifies some universally valid relationship between certain *kinds* of conditions and certain *kinds* of events. Such kinds are simply classified observables. In part, a general law - for example, "All immersed objects having a density less than that of the surrounding liquid necessarily float" - is a tool for classifying observables into groups or kinds. Spatio-temporally unique circumstances are rendered familiar, intelligible, through classification. In particular, classification using general laws draws attention to a necessary *relationship* between the classified observables. The example defines a *causal* relationship between the antecedent conditions, 'an immersed object having a density less than that of the surrounding liquid', and the consequent event, 'an object floating'. In other words, a set of certain kinds of conditions sufficient for one kind of event is specified by such a law. Every event is presumed to have a cause, and it is no accident that the majority of laws of a mature science are causal. Explanations that incorporate causal laws are *causal explanations*, under the D-N model.

Satisfying the first formal requirement goes a long way towards meeting the others. Logical deduction of the *explanandum* is possible with a causal law among the statements of the *explanans*. Yet what is deduced is a statement about the occurrence of a particular fact (*explanandum*). Although the laws are about kinds of observables, a scientific explanation is typically about *particulars*. The aboutness relationship is an informal condition of scientific explanation. Thus statements of initial and boundary conditions must accompany relevant laws if the *explanandum* is to be *deduced* from the

explanans. Having incorporated the example law into an *explanans*, to causally explain the occurrence of a particular floating object p, the *explanans* further requires statements of initial and boundary conditions, such as:

8(a) At time t, object p is placed into a beaker filled with liquid q.

(b) Object p has a density of x grams per cubic centimeter.

(c) Liquid q has a density of y grams per cubic centimeter, where y is greater than x.

Coupling statements 8 (a-c) with the example law will entail the *explanandum*: "Therefore, at time t, object p floats in liquid q."

Third and last, an explanation is empirically meaningful if what it purports about the world is either true or false and, if this can be ascertained through direct or indirect observation. Statements of the initial and boundary conditions are just reports from the observer's notebook. Although scientific observation is an acquired skill, the empirical meaning of the observation statements is relatively unproblematic. Provided an observation statement is an honest report of what is at hand in a given experiment, that is, provided no reference is made to immeasurable entities, the statements will be either true or false. Further, in large part the *explanandum* too is an observation statement. Yet the *explanandum* also involves non-empirical, logical terms for example, "therefore", "hence" or "thus". Of course, given the statements about initial and boundary conditions, the use of such non-empirical terms is licensed by the general law incorporated into the *explanans*. Thus the problem of the empirical meaning of any scientific explanation revolves around the problem of ascertaining the empirical status of the relevant general law(s) used in the *explanans*.

General laws do not refer directly to particular observables, but to classes; their reference to particulars is roving and oblique. Hempel has written an excellent chapter on the problems and changes in the Empiricist criterion of cognitive significance (*Aspects*, 101-122). In essence, the Herculean task is to specify the requirements for a scientific language that admits of logical terminology, extralogical universal and existential statements, yet excludes the use of metaphysical and theological statements which partly refer to observables and partly to that which is unobservable. Extralogical statements are admissible if they are *testable*.

This brings our very brief discussion on Hempel's theory of D-N explanation to his thesis that explanation and *prediction* are *structurally isomorphic*. Explanation and prediction have the same logical structure according to Hempel, such that any D-N explanation could have served to predict the occurrence of the event otherwise explained. In otherwords, the *explanandum* can operate as either a statement of that which has occurred and requires explanation, or as a statement of that which is expected to occur and wants prediction. The only difference between explanation and prediction, as hereconstrued, is the time at which the observer constructs a D-N argument. Explanation is either during or after the fact, and prediction is accomplished before the fact.

The thesis is most powerful, for it provides the means to test purported explanations. An explanation is empirically meaningful if it could have served to predict beforehand what it purportedly explains. The structural isomorphism of explanation and prediction is attributable to the universality of a scientific law, and thus by

means of general laws explanations are testable and their empirical meaning is ascertainable. Under the D-N model, the use of general laws is of paramount importance if the occurrence of a natural event is to be explained with the precision required of formal understanding.

The D-N model has been criticized as factitious. Few examples of explanations that are considered scientific, it is argued, evince the structural characteristics of a deductive-nomological explanation. Its structure is unusual in physics and chemistry, rarer still in the biomedical sciences, unheard of in psychology and sociology and, some say, even impossible or wrong-headed in history. The most characteristic feature of D-N explanation is at once its greatest point of contention: that scientific explanation requires the subsumption of its subject matter under general regularities or laws.

Hempel's model is not, however, simply a description of actual explanations offered in empirical research. Thus, even though an actual explanation may not have the appearance of a deductive argument, *in principle*, it must be able to conform to the D-N model to count as a scientific explanation. Moreover, explanations that are generally recognized as scientific are able to conform with the D-N standards. And perhaps of greatest value, the model is sufficiently clear to permit problems to be identified within recalcitrant explanations. If problems can be identified one has the direction toward a solution. Finally, with explanation as the conceptual means to scientific knowledge, the D-N model becomes a convenient gauge of scientific maturity, and formal understanding.

3.4 FUNCTIONAL EXPLANATION AND THE D-N MODEL

With the above position in mind, the D-N model provided the

background to Hempel's study of functional explanation. As he perceives things, functional explanation is invoked to explain the recurrent activity of systems, such as biological systems. To evaluate the logic of functional explanation, Hempel considers the heartbeat example:

> (1) "The heartbeat in vertebrates has the function of circulating blood through the organism." (305)

The heart of course is an organ of the circulatory system, a system characterized by the recurrent activity of having blood surge through its vessels. The question arises as to the meaning of the term "function". Clearly, the term is not equivalent to "effect", since not every effect of the heart(beat), nor for that matter of any other biological item (organ, structure, process, etc.), would be considered a function. Heart sounds are recurrently effected by the heartbeat, yet these sounds have no obvious function in the system that produces them. Likewise, the heart adds to the total body weight, but such an effect is not a function of the heart. Having an evident need for a qualification, Hempel suggests that something counts as a function if it is an *indispensable effect*. Sentence (1) thus becomes:

> (1) "The heartbeat has the effect of circulating the blood and this ensures the satisfaction of certain conditions (supply of nutriment and removal of waste) which are necessary for the proper working order of the organism." (305)

Hempel thus construes the function of a given trait in terms of its causal relevance to the satisfaction of certain necessary conditions of proper working order. As we soon will see, Cummins argues that in specifying the notion of function by reference to

the concept of proper working order Hempel falls victim to a trilemma.

But to Hempel, the most troublesome features of functional explanation arise if explanations having the form of '1' are further elaborated to fit the logical structure of the D-N model. He writes:

> "Suppose, then, that we are interested in explaining the occurrence of trait i in a system s (at a certain time t), and that the following functional explanation is offered:
>
> (1'a) At t, s functions adequately in a setting of kind c (characterized by specific internal and external conditions).
>
> (b) S functions adequately in a setting of kind c only if a certain necessary condition n is satisfied.
>
> (c) If trait i were present in s then, as an effect, condition n would be satisfied.
>
> (d) (Hence), at t, trait i is present in s." (310)

Note that statements (a-c) constitute the *explanans*; d is the *explanandum*. Both b and c are law-like; a is a statement of the initial and boundary conditions, and d is a statement of the occurrence of the particular trait i. A very pointed criticism advanced by Cummins against the above scheme, which we will detail later, is that it pictures functional explanation as a deduction of the *presence* of some trait i from the *function* of that trait: Functional explanation is reasoning from effect to cause, not cause to effect, as in usual causal explanation.

However, the trouble that receives Hempel's attention is that 1" affirms the consequent: d does not follow logically from a, b, and c. At the very outset, functional explanation fails to qualify as a scientific explanation because it does not meet the second requirement of the D-N model: deducibility. Hempel's attempt to

resolve this trouble finds functional explanation skewered at the points of a trilemma.

The trilemma may be developed with the use of a homely example:

(9a) Yesterday, Imacow (system s) functioned adequately on farmer Brown's farm.

(b) Imacow functions adequately on farmer Brown's farm only if she is milked and fed twice daily.

(c) If farmer Brown were present then, as an effect, Imacow would have been milked and fed twice yesterday.

(d) (Hence), yesterday, farmer Brown was present and twice milked and fed Imacow.

Clearly, the argument is invalid. The milkmaid or son Robert may have in fact milked and fed Imacow. To salvage the validity of 9, c must be altered to read:

(c') If and only if farmer Brown was present then, as an effect, Imacow would have been milked and fed twice yesterday.

Farmer Brown may be inclined to argue that Imacow functions adequately only when he and only he feeds and milks her: she produces more milk and feeds better, and therefore functions (more) adequately. It is, however, difficult and perhaps impossible, to support such a claim with appropriate evidence. Imacow might change her finicky habits in the course of testing the hypothesis of farmer Brown's *functional indispensability*. Should this occur, farmer Brown may further contend that after the testing Imacow is not the same cow, for now she functions adequately in the hands of anyone who milks and feeds her. However, as Hempel correctly points out:

> "... this type of argument would safeguard the postulate of the functional indispensability of ... (farmer

> Brown) ... against any conceivable disconfirmation but at the cost of turning it from an empirical hypothesis to a covert definitional truth." (312)

In short, the *explananda* of 1" and 9 can be deduced from their respective *explanantia* only at the cost of transforming the purported explanations into necessary truths. Emptied in this way of all empirical content, functional explanation fails to satisfy the third criterion of D-N explanation.

One might argue against the above conclusion that, although farmer Brown is dispensable, the functional items *within* a biological system are not functionally indispensable. For example, Imacow's heart is indispensable to her functioning adequately. Yet it is conceivable that mechanical hearts could some day replace normal biological ones. Moreover, even biologists entertain such a "principle of multiple solutions", as Hempel has seen:

> "This principle, which has been emphasized by functionally oriented biologists, states that for a given functional problem (such as the perception of light) there are usually a variety of possible solutions, and many of these are actually used by different - and often very closely related - groups of organisms." (311)

Biological items are as functionally dispensable as farmer Brown.

Thus far Hempel's analysis of functional explanation leads to the horns of a dilemma: on one hand, if we attempt to satisfy the empirical significance criterion of D-N explanation, we sacrifice deducibility; on the other, if deducibility is restored, empirical meaning must be abandoned. Empirical meaning and deducibility would appear to be mutually exclusive properties of functional explanation.

To try to avoid the dilemma, Hempel suggests that premise

1"c be adjusted further to read:

> (c") "I is a class of all empirically sufficient conditions for the fulfillment of the requirement of n, in the context determined by system s in setting c." (312)

At best, however, entering premises 1"a, b, and c" into an *explanans* enables the deduction of:(1"d") "Some one of the items included in class I is present in system s at time t." (312) In addition, this "weak" explanation is valid only with the further qualification that class I is not empty. In terms of our homely example, the qualification licenses the deduction that, "Yesterday, *someone* twice milked and fed Imacow." Hempel is not convinced that such a qualification offers a safe retreat from the dilemma, for it leads to the further difficulty of failing to satisfy his informal criterion of *specificity*. Explanation in science is considered to explain the occurrence of particulars, yet functional explanation, if valid and empirically meaningful, cannot explain why one state of affairs is necessarily present, rather than an alternative. This embarrassment is the third point of the trilemma of which I spoke.

Hempel offers the following stern conclusion on functional explanation:

> "... its explanatory force is rather limited, in particular, it does not provide an explanation of why a paricular item i rather than some functional equivalent of it occurs in system s." (324)

No doubt the severe blow thus dealt to functional explanation prompted Cummins to examine the logical underpinnings of Hempel's position.

3.5 CUMMINS' NEGATIVE THESIS

Cummins' primary criticisms of Hempel's theory of functional

explanation have been mentioned in passing. The first concerns what counts as a function, and the second concerns what is deduced in functional explanation.

A function is an indispensable effect of some item i, Hempel contends, although the item itself may not be indispensable in the sense that an alternative item could theoretically produce the same indispensable effects, or functions. Such effects are indispensable to, as yet, unspecified conditions of proper working order of a given organism. Hempel's notion of function is what Cummins has called a "selected-effects" theory.

The idea of a function being an effect indispensable to the proper working order of an organism is circular, Cummins argues.

> "... it seems clear that for something to be in... adequate or effective or proper working order is just for it to be capable of performing its functions adequately or effectively or properly." (753)

What Cummins construes to be tautological becomes the third point of a trilemma, the development of which begins with the following argument:

> "... if we identify the function of something x with those effects of x which contribute to the performance of some activity a or the maintenance of some condition c of a containing system s, then we must be prepared to say as well that a function of s is to perform a or to maintain c." (753)

The argument is a logical howler, and I will postpone further comment. For Cummins, however, Hempel's analysis of function either launches a regress, or "breaks down at some level for lack of functions." (753-54) As Cummins views things, Hempel's definition of "function" places the burden of meaning upon the unspecified meaning of the next 'higher order' function, to which function the antecedent function contributes, and so on up the organizational 'hierarchy'.

Or else, having reached the final step of the hierarchy, there are no more functions to which reference can be made and the 'analysis' of a system's functions breaks down. Thus, for example, "Perhaps protozoans have no functions." (753) If a system has no function, the antecedent functions which contribute to the proper working order of the system-without-a-function cannot have their meanings specified, and the regress proves to be fruitless. Should the containing system have a function, Cummins would appear to be unhappy about the need to regress to that function to understand the meaning of the contained functions. An adequate definition of "function" should provide an understanding why, for example, *blood circulation* is the heart's function, without reference to a containing system's role in an economy, natural or otherwise.

To avoid what is so far a dilemma, Cummins says it must be determined:

> "...whether a thing's function can plausibly be identified with those of its effects contributing to the production of some activity of, or maintenance of some condition of, a containing system, where performance of the activity in question is not a function of the containing system." (754)

A possible solution Cummins entertains specifies the notion of proper working order in terms of *species survival*. On this criterion, a function is viewed to be an effect of an item that either directly or indirectly contributes to the organism's reproductive potential, and hence to the survival of its species. The elimination of excess water, for example, satisfies this criterion, since the activity keeps fresh-water protozoa from bursting with extreme osmotic pressures, and this helps to maintain the reproductive potential of such organisms.

Although many applications of the notion of function satisfy the criterion, Cummins finds it too restrictive even within the field of biology. Should fresh-water protozoa be successfully introduced into salt-water, Cummins speculates, their contractile vacuoles could not effect an elimination of an excess of water, for water is scarce in a salt-water environment, yet the potential for the activity remains and it is thus correct to refer to that as the function of the contractile vacuole. Again, flight is the function of the wings of a pigeon, yet flight may simply waste energy in an environment free from predators and abundant in food. Flight then has no survival value, yet it remains the function of a pigeon's wings. Only if an item ceases to have the *capacity* for an effect is it proper not to select it as a function of the item, Cummins argues. Such is the case with the wings of ostriches and penguins, which do not have the capacity for flight, and thus flight is not considered their function. Cummins' positive thesis develops the idea that the notion of function means having a capacity.

However, these examples tell Cummins that the concept of a biological function must be specified "independently of evolutionary considerations." Moreover, in the restricted cases where such considerations are relevant, no philosophic analysis is provided to explain why using the term "function" is appropriate. What is offered instead is just a description of actual usage. To summarize his criticism of Hempel's notion of function, Cummins writes that a "selected-effects" theory either:

> "...launches a regress, or terminates in an unanalyzed functional ascription; if not ... then it leaves open the very question at issue, viz., why are the selected effects seen as functions?" (757)

On the matter of what counts as a function, Cummins thinks he has Hempel's theory skewered on the points of a trilemma.

Cummins' criticism of Hempel's thinking on what is deduced in functional explanation is straightforward and effective. Hempel assumes that a functional explanation is an argument which leads from the observation of certain "selected-effects" to the deduction of the presence of a particular item having those effects. Hence functional explanation is used to deduce hearts, kidneys, pancreases, and so on, from the recurrent activities or effects of blood circulation, waste removal, and blood-sugar regulation, respectively. Clearly, however, to explain the presence of any such organ by appealing to what the organ does "... is to 'explain' its presence by appeal to factors that are causally irrelevant to its presence." (745) Hearts, kidneys and pancreases exist in vertebrates not because they do what they do, but because the organism's genes specify the development of such functionally significant items. Explanations of why something exists which make reference to the thing's function are appropriate only when explaining the functional significance of human artifacts. A reference to the function of such an item explains "why the thing was put there." To explain, for example, why a door stop is wedged against a door, legitimate reference should be made to the function of the door stop:'to hold open the door'. Cummins concluded:

> "The validity of the sort of explanation in question should depend upon the assumption that the thing functionally characterized is there as the result of a deliberate action." (747)

Since this assumption is not available to the study of biological systems, functional explanation is inappropriate in biology.

These then are Cummins' main arguments against Hempel's theory of functional analysis. Functions are not selected effects, and functional explanation is not a deduction from the occurrence of some functional activity to the presence of some item functionally characterized. Critical examination will follow my discussion of Cummins' alternative theory of functional analysis which is meant to overcome the problems he found with Hempel's theory. Cummins presents a position on what counts as a function, and then sketches a model of explanation, termed "function-analytical explanation", which is supposedly both distinct from D-N structured functional explanation and such as to legitimize the use of functional terminology in biology.

3.6 CUMMINS' POSITIVE THESIS

In Cummins' view, to name a function is to name a capacity of some system. This is qualified with the understanding that a system may be capable of a function f, but never realize f. On the other hand, i may function as an f, although f may not be the function of i. Thus a pump is capable of pumping, and pumping is its function, yet circumstances may be such that it never actually pumps. It may for instance corrode away in the corner of a warehouse full of pumps. In contrast, another item may function as a pump, a surgeon's hand for example, although pumping may not be its specific function. However this may be, according to Cummins, if something admits of a certain function, or if it realizes its proper function, it must obviously have the capacity for that function. In short, functions are capacities.

A capacity is analyzed into a *disposition* to behave in a predictable manner, given a set of antecedent "precipitating conditions" that activate the disposition in question. The relation between

the precipitating conditions and the dispositional act is a law-like regularity referred to by Cummins as a "dispositional regularity". The capacities of a system are thus its dispositional regularities, and when a system manifests a specific disposition, one of its functions is said to operate. Thus the heart functions as a pump, since pumping is a disposition that is regularly manifested.

Functional analysis is motivated by a type of problem in formal understanding, on Cummins' account. Explaining how the total capacities of a given organism are possible, or explaining how any one capacity can operate, is the type of problem a functional analysis is invoked to explain. He contends that functional *explanation* has matters reversed; one cannot explain the presence of an item or cause by reference to its capacity or effect, rather, functional *analysis* should be designed as explanations of various capacities by reference to the items that cause the capacities to be manifested. That is, functional analysis should be designed to explain "how the organism comes to exhibit certain characteristics of behaviour." (749) To explain the function of the heart in vertebrates is to indicate how the capacity for circulation is possible.

So construed, how does the logic of explanations rendered in functional analysis differ from that of the ordinary causal explanations we looked at earlier? The difference lies in what Cummins calls the *style* of function-analytical explanation. Such explanations are not deductive-nomological, or so he claims.

Cummins characterizes two "styles" of explanation, the "subsumption strategy" and the "analytical strategy", and ventures that "the appropriateness of function-ascribing statements corresponds

to the second of these two strategies." (758) The subsumption strategy is appropriate whenever a general law can be used to explain the relation between certain kinds of (antecedent) precipitating conditions and certain kinds of dispositional regularities. Though not unique to a given dispositional regularity, the law points out the kind of precipitating conditions that, if present, are sufficient to the disposition's being regularly manifested. The precipitating conditions and the dispositional regularity are simply subsumed under a general law, and presumably, explanations that engage the subsumption strategy enjoy the style (logic) of D-N explanation. No examples are drawn from biology for this style of explanation. However, the suggestion is made that if an analysis can reach "the level of pure physiology" the subsumption strategy becomes appropriate. Yet, again, it is unclear what "pure physiology" means - biochemistry? This obscurity weakens the distinction between strategies.

Leaving the trouble aside for the moment, Cummins argues that if the causes or precipitating conditions *are* unique to their effects or dispositional regularities, the analytical strategy is the one to use for explanation. He writes:

> "Rather than subsume a dispositional regularity under a law not special to the disposed objects, the analytical strategy proceeds by analyzing a disposition d of a into a number of other dispositions $d_1 \ldots d_n$ had by a or components of a such that programmed manifestation of the d_i results in or amounts to a manifestation of d." (759)

A complex capacity (disposition or function) is thus explained by analyzing it into component capacities. In other words, the analytical strategy proceeds by breaking down a macroevent d into a number of microevents $d_1 \ldots d_n$ which, when manifested in sequence,

or according to a "program", emerge as the macroevent d. Moreover:

> "Such an analysis allows us to explain how the device as a whole exercises the analyzed capacity, for it allows us to see exercises of the analyzed capacity as programmed exercise of the analyzing capacities. In this case, the "program" is given by the lines indicating how the components are hooked up. (Of course, the lines themselves are function symbols.)
>
> Functional analysis in biology is essentially similar. The biologically significant capacities of an entire organism are explained by analyzing the organism into a number of "systems" - the circulatory system, the digestive system, the nervous system, etc., - each of which has its characteristic capacities. These capacities are in turn analyzed into capacities of component organs and structures. Ideally, this strategy is pressed until pure physiology takes over, i.e., until the analyzing capacities are amenable to the subsumption strategy. We can easily imagine biologists expressing their analyses in a form analogous to the schematic diagrams of electrical engineering, with special symbols for pumps, pipes, filters, and so on. Indeed, analyses of even simple cognititive capacities are typically expressed in flow charts or programs, forms designed specifically to represent analyses of information processing capacities generally." (760-61)

Presumably, schematic diagrams are not the only method to express such an analysis. The causal relation between the antecedent microevents and the resultant macroevent is expressible in a narrative which would describe a continuous series of microevents that produces a programmed manifestation of the macroevent. Such a narrative could be supplemented with a schematic flow chart or diagram that would picture the complex causal network described in the narrative. Any physiology textbook abounds in examples of flow charts and descriptive narratives.

Having outlined the explanatory style of the analytical strategy, Cummins returns to the problem of the appropriate use of

the term "function" in biology. His views on the notion of function and on the analytical strategy dovetail into his theory of "function-analytical explanation". According to Cummins, it is only appropriate to use functional phraseology within the context of some analysis of the capacities of a containing system. Without a containing system, such an analysis is impossible, and without such an analysis, functional terminology is inappropriate.

> "It is appropriate to say that the heart functions as a pump against the background of an analysis of the circulatory system's capacity to transport food, oxygen, wastes and so on, which appeals to the fact that the heart is capable of pumping." (762)

Function-ascribing statements are therefore implicitly dependent upon an analytical context. The general idea is thus defined:

> "x functions as a ø in s (or: the function of x in s is to ø) relative to an analytical account A of s's capacity to ψ just in case x is capable of ø-ing in s and A appropriately and adequately accounts for s's capacity to ψ by, in part, appealing to the capacity of x to ø in s." (762)

Moreover, the above definition of appropriate use of functional phraseology,

> "... makes no provision for speaking of the function of an organism except against a background analysis of a containing system (the hive, the corporation, the ecosystem). Once we see that functions are appealed to in explaining the capacities of containing systems, and indeed that it is the applicability of a certain strategy for explaining these capacities that makes talk of functions appropriate, we see immediately why we do not speak of the functions of uncontained containers." (763)

That functions explain the capacities of containing systems is a point Cummins repeated throughout his criticism of Hempel's theory of functional explanation. The companion claim is now made

that capacities are so explained by virtue of a "function-analytical explanation" which "mediates the connection between functional ascriptions (x functions as a ø, the function of x is to ø) and the capacities of the containers." (763) Functions are linked to capacities, or equally, "capacities emerge as functions", within the context of a function-analytical explanation. A function-analytical explanation is just an application of the analytical strategy that expressly uses function-ascribing statements.

3.7 CRITIQUE OF CUMMINS' THESES

If we work back from Cummins' theory of function-analytical explanation, to his two criticisms of Hempel's position, we find fatal problems with both his positive and negative theses. In the positive thesis, Cummins' distinction between the analytical and subsumption strategies appears to rest upon a misunderstanding of the nature of an empirical law, and of its role in causal explanation.

The subsumption strategy, we saw Cummins remark, employs "laws governing the behaviour of things generally, not just things having" the disposition in question. How is the phrase "governing the behaviour of things generally" to be understood? Judging from the manner in which Cummins uses it, it appears to mean "governing things in general", for the laws cannot provide an account that is "unique" to what has occurred, but can only provide non-unique accounts. Examples of non-unique laws that govern the behaviour of things in general must be:

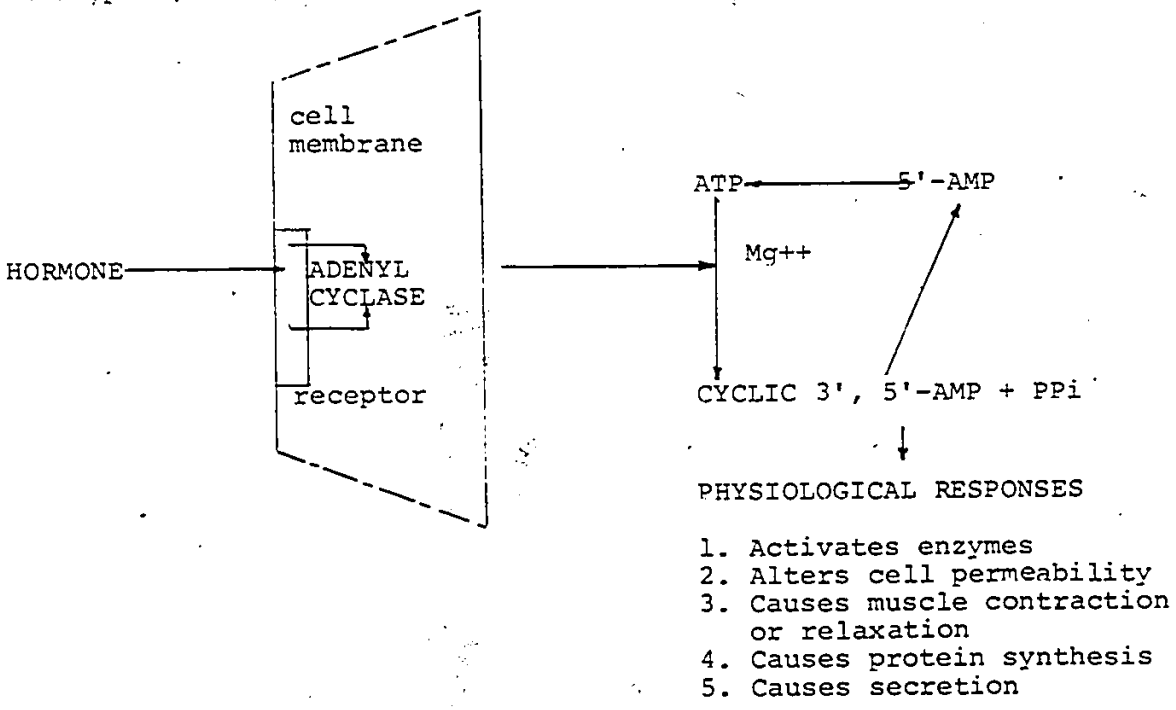
i) Every event has a cause.
ii) Two objects cannot occupy the same point in space at the same point in time.
iii) An object's spatio-temporal framework is anisotropic.

Yet these are not empirical laws, but metaphysical statements referring to the basic structure of the world, the basic structure of things in general. In sharp contrast, empirical laws do refer to *specific* sorts of things. They indicate of a particular occurrence that its presence is not unique, but an instance of a familiar causal routine. Cummins has fallaciously reasoned that because a law denied uniqueness to a particular event, the law itself becomes non-unique and a governor of things in general. However, the denial of uniqueness is specific to certain classes of particulars; empirical laws are not metaphysical statements, and thus the former have explanatory efficacy within attempts to understand particular occurrences. Archimedes' principle, to which Cummins refers in his example of explanation via the subsumption strategy, is not an example of a law that refers to things in general; Archimedes' principle refers *only* to those objects evincing the behavioural characteristic of 'elevancy' when immersed into a liquid of greater density. Empirical laws do not govern the behaviour of things generally, only the general (i.e. routine) behaviour of specific sorts of things.

Supposedly, explanation via the analytical strategy does not rely on general laws, but proceeds by breaking down a complex event or system into components. Yet how *could* the analytical strategy proceed without general laws to classify and explain various dispositional *regularities*? How, for example, could it be determined that the set $d_1 \ldots d_n$ is the actual cause of d, and not some other set $c_1 \ldots c_n$, unless a general law is available which specifies that the former is the kind of thing that can give rise to d, whereas the latter is not? Moreover, it does not appear possible to refer to a dispositional regularity as a regularity *per se* without supposing

that it is an instance of a causal routine, 'governed' by an empirical law. No number of observation statements, such as $d_1 \ldots d_n$ and d, can do more than hint at a statistical correlation between the former and the latter. Any 'if-then' proposition implies reference to a general law.

Moreover, flow charts are not used in biology to depict unique events, and they *do* imply 'if-then' propositions. In fact, a flow chart is useful precisely because it is a kind of *explanation-sketch* of a causal routine. In the biosciences, the lines of a flow chart typically represent causal routines and go proxy for general hypotheses about a given physiological mechanism. Let us consider an example of a very useful theory of physiology: the theory explaining the cyclic-AMP mechanism by means of which *many* hormones exert their control over cell function. The example is from Guyton's *Textbook of Medical Physiology* (W.B. Saunders, Philadelphia, 1976).



The following narrative accompanies the flow chart:

> "Many hormones exert their effects on cells by first causing the substance *cyclic* 3', 5' *adenosine monophosphate* (cyclic AMP) to be formed in the cell. Once formed, the cyclic AMP causes the hormonal effects inside the cell. Thus, cyclic AMP is an intracellular hormonal mediator. It is also frequently called the *second messenger* for hormone mediation - the "first messenger" being the original stimulating hormone.
>
> (The flow chart) illustrates the function of cyclic AMP in more detail. It is believed that the stimulating hormone acts at the membrane of the target cell, presumably combining with a specific receptor for that particular type of hormone. The specificity of the receptor determines which hormone will affect the target cell. After binding with the receptor, the combination of hormone and receptor activates the enzyme *adenyl cyclase* in the membrane, and the portion of the adenyl cyclase that is exposed to the cytoplasm causes immediate conversion of cytoplasmic ATP into cyclic AMP. The cyclic AMP then initiates any number of cellular functions before it is destroyed - functions such as activating enzymes in the cell, altering the cell permeability, initiating the synthesis of specific intracellular proteins, causing muscle contraction or relaxation, initiating secretion, and many other possible effects. The types of effects that will occur inside the cell are determined by the cell itself. Thus, a thyroid cell stimulated by cyclic AMP forms thyroid hormones, whereas an adrenocortical cell forms adrenocortical hormones. On the other hand, cyclic AMP affects epithelial cells of the renal tubules by increasing their permeability to water." (990-91)

Although the above explanation appears less like a D-N explanation and more like a narrative about a continuous series of events, as per the analytical strategy, it abounds in tacit references to general laws. The following is one law implied in the narrative:

> (L_1) Cyclic AMP is a certain *kind* of hormone, which is a certain *kind* of chemical substance, and if, at time t, cyclic AMP is present within a certain *kind* of membrane, cyclic AMP is disposed to effect certain *kinds* of changes within the cell proper.

L_1 is a general law insofar as it asserts a *causal routine* between certain kinds of conditions and events: respectively those which include, among other conditions, the presence of a cell membrane and cyclic AMP molecules, and the event of a certain kind of physiological response (e.g., protein synthesis). L_1 does not refer to particular molecules of cyclic AMP, to particular cell membranes, nor to individual examples of protein synthesis. If the example from Guyton captures what Cummins had in mind about the analytical strategy, there appears to be nothing about the strategy to prevent it from collapsing into the subsumption strategy - i.e., into what Hempel termed a D-N explanation. Function-analytical explanation, of which the Guyton narrative is an example since it uses functional phraseology within the context of an analysis of the cyclic AMP mechanism, fails to secure a logic (style) distinct from D-N explanation.

It was remarked earlier that Cummins rejects functional explanation as inappropriate to accounts of natural systems, because such accounts cannot presume that the functionally characterized items exist as the result of a deliberate act. I also said the criticism is effective against Hempel since Hempel presumes at the outset that one ought to be able to deduce the presence of *particular* items from their functional effects or capacities. As effective as it is, Cummins inconsistently permits functional explanation, as arguments from effects to specific causes, to enter his own theory of function-analytical explanation. On page 751 he claims: "the function of an organ is appealed to to explain the biological capacities of the organism containing it." Functions *explain*, they are not simply appealed to in descriptions. What is the logic of such

explanations which show how it is that an organism has certain capacities because of the presence of certain organs? What logical rule allows the inference? Cummins' response is most peculiar:

> "Our best (only) explanation of photosynthesis requires chlorophyll, and our best explanation of coordinated activity requires nerve tissue. But once we see what makes this reasoning legitimate, we see immediately that inference *to* an explanation has been mistaken for an explanation itself." (748)

The sort of explanation inferred when we "legitimately" reason from function to cause is, of course, a function-analytical explanation. In other words, we are to believe that reasoning from function to cause is legitimate since it is an inference to some causal explanation of the form, $d_1 \ldots d_n \longrightarrow d$, or $i \longrightarrow f$. But this appears to affirm the consequent, i.e., he appears to argue that '$f \longrightarrow i$' is legitimate, because $(f \longrightarrow i) \longrightarrow (i \longrightarrow f)$ and '$i \longrightarrow f$' is true! If Cummins has not fallen prey to this fallacy, I am uncertain what he means; he certainly offers no philosophical analysis of why inferences from f to i are legitimate, he simply uses them. Yet this is precisely the problem Hempel tackles in his essay: specifying the logical grounds that justify the prevalent inferences in biology from functional capacities to the presence of items functionally characterized. In short, Cummins' positive thesis not only fails to provide a convincing alternative to the D-N model, but returns us to Hempel's problem of attempting to fit functional explanation into the deductive-nomological framework.

Cummins' negative thesis attempted to develop a trilemma

from Hempel's apparent equation of the notions of function and proper working order. The first and second points of the trilemma, however, are developed from yet another fallacy, and the third omits to consider Hempel's discussion of 'self-regulation', which is a non-tautological account of the notion of function. Cummins argued that Hempel's theory of functions either launches a regress or breaks down at some level for want of functions in the organizational hierarchy. This is true provided we accept the assertion:

> "... if we identify the function of something x with those effects of x which contribute to the performance of some activity a or to the maintenance of some condition c of a containing system s, *then we must be prepared to say as well that a function of s is to perform a or to maintain* c." (753, underlining mine)

Had he applied the argument to an example, he would have questioned its validity. Although one is prepared to claim, 'the function of the contractile vacuole in fresh-water protozoa is the elimination of excess water from the organism', one is not inclined to further claim, 'therefore the function of a protozoan is the elimination of excess water from itself'. What is true of a part is not necessarily true of the whole. This is entirely reminiscent of the old argument, "Since each part of the universe has a function, the universe itself must have a function", which is now a textbook example of the informal fallacy of *composition*. The initial points of Cummins' trilemma are thus developed from an application of the fallacy of composition!

The third point dissolves as well, rendering the trilemma completely innocuous. Cummins failed to consider Hempel's argument that functional language is meaningful only if "it is properly

relativized" to an adequate "general hypothesis of self-regulation". Hempel's notion of function is not a definitional truth; in fact, an entire section of his essay is devoted to the problem of "the empirical import of functionalist terms and hypotheses". (319-25) Here, he elaborates the meaning of "the key terms of functional analysis" such as, "function", "need", "adequate or proper working order or functioning", "adaptation" and "adjustment". Two definitions of "function" are examined and may be generalized as follows: a function is that selected effect of an item which,

a) satisfies a need (i.e., a set of conditions individually necessary and jointly sufficient to the organism's *survival*).

b) is capable of adjustment or adaptation within a system.

The concept of survival (used in a) does not refer to the conditions "of just the barest survival but as conditions of persistence in, or return to, a "normal", or "healthy" state, or to a state in which the system is a "properly functioning whole"." (321) A function is a selected effect which satisfies a need, which need is some condition necessary to survival, where survival is the maintenance of what Nagel termed a G-state. Thus, for example, controlling the metabolic activity of cells is said to be the function of thyroxine, in Hempel's sense, because this is necessary to maintaining, among other things, the G-state of thermal homeostasis. Equally, the concepts of adaptation and adjustment (used in b) are meaningful provided they are specified in terms of "some standard; otherwise they have no definite meaning and are in danger of being used tautologically or else subjectively, with valuational overtones." (322) The standard of normality or adjustment varies

between groups of organisms, and an appropriate functional analysis must therefore be specific with respect to the scope of application to be empirically meaningful. He writes:

> "In the functional study of a given system s, the standard would be indicated by specifying a certain class or range R of possible states of s, with the understanding that s is to be considered as "surviving in proper working order", or as "adjusting properly under changing conditions" just in case s remains in, or upon disturbance returns to some state within the range R. A need, or functional requirement, of system s relative to R is then a necessary condition for the system's remaining in, or returning to state in R; and the function, relative to R, of an item i in s consists in i's effecting the satisfaction of some such functional requirement." (323)

In other words, a function is a selected effect that satisfies a functional requirement. A functional requirement is a particular condition (e.g., blood-circulation or vasodilation) necessary to the system's both remaining in and upon disturbance returning to a specifiable set of possible G-states of the system. Functional requirements are conditions necessary to a system's capacity to regulate itself or maintain its normal biological activity. Thus a function (e.g., blood-circulation) is a condition without which a system could not maintain its normal biological activity; it is a necessary condition of self-regulation. What guarantees the empirical meaning of Hempel's definition of function is the specificity of the corresponding general hypothesis of self-regulation. The hypothesis should be species specific, which enables it to be tested against the observed behaviour of a group of organisms. And testability provides for empirical meaning.

Returning for a moment to the fate of Cummins' positive thesis, function -analytical explanation does not differ, in

principle, from deductive-nomological explanation. Cummins fails to secure a distinct logic true for functional explanation, contrary to his promise. Second, his analysis of 'function' into 'capacity', and the latter into 'the dispositional behaviours of the parts of the system', asserts no content which contradicts Hempel's selected-effects theory of function. The analyses are compatible, and merely express the same idea in different terminology. For instance, the idea of, 'The function of the heartbeat is blood circulation' is captured equally well in either:

A capacity of the heartbeat	which helps to maintain a G-state	is blood circulation
A selected-effect of the heartbeat		

The strong point of Cummins' essay is, as stated previously, his criticism of Hempel's assumption that functional explanation ought to be a deduction of the presence of an individual item from the observation of certain biological activities. His criticism is effective, but in developing his positive thesis on the analytical strategy he throws out the baby with the bath-water. Hempel is wrong about what is deduced in a functional explanation, but more importantly, he is correct about the deduction thesis: that it is necessary to explanation, functional or otherwise. Furthermore, Hempel's trilemma on functional explanation can be avoided.

CHAPTER II.

SHERRINGTON'S DISCOVERY OF INTEGRATIVE ACTION

"The purpose of a reflex seems as legitimate and urgent an object of natural inquiry as the purpose of the colouring of an insect or a blossom."

-- Sherrington, 1906

The Integrative Action of the Nervous System

1. PRELIMINARY STATEMENT

The defence for my thesis that teleology is necessary to understanding in physiology is grounded in Sherrington's theory of the integrative action of the nervous system. In the present chapter, I discuss how the theory develops. 'Integrative action' is a causo-physical theory that explained, for the first time, the striking purposiveness and coordination long observed in spinal reflex responses. It would also be the conceptual framework for later studies of the brain, as I have said, and would become as well the general model for purposive behaviour in physiology. Though causo-physical, the integrative action theory is also teleological, for, in Sherrington's view, we don't understand a given reflex until we know its aim or purpose. Understanding the spinal reflex was of course Sherrington's life-time ambition and occupation. As we shall see, through successively more complex orders of coordination, the "meaning" or purpose of a particular reflex act becomes more certain. This is not to suggest that the assignment of a purpose to a given reflex is an arbitrary, subjective imposition by the observer himself. On the contrary, assigning purposes to reflex acts is based on the particular characteristics which can be directly observed in the coordinated responses themselves. Sherrington's method of assignment is therefore empirical, and not unlike the method of generating (other) empirical hypotheses. I shall discuss his rules for the assignment of purpose.

The present discussion of integrative action will thus provide the material for demonstrating in Chapter III how Sherrington's theory of biological organization is superior to Nagel's. The reason I believe

Sherrington's theory is successful, whereas Nagel's is not, is precisely because the former develops the use of teleology, whereas Nagel's tries to avoid it. My thinking is that, if we attempt to thoroughly understand one area of physiology which deals with purposive behaviour, we will be able to better appreciate the pitfalls in Nagel's position. 'Integrative action' is thus a case study of a scientific theory on purposiveness, concerning which I hope to draw correct conclusions about both issues in the debate on teleology.

During the course of my discussion on integrative action, I will, in passing, bring attention to Sherrington's method of research and explanation, which will prove useful for my critique in Chapter IV, on the logic of functional explanation, as I see it.

2. INTRODUCTION

Every schoolboy knows the knee jerk is a reflex. It is perhaps the only reflex known to most people. The commonplace status of this reflex is surprising when it is learned that less than two generations ago, the knee jerk was the one neuromuscular response that leading physiologists of the day could not classify as being reflex. The reason was that the knee jerk appears to have no *latent period*, the period between the moment a stimulus is applied and the moment a response becomes evident. The latent period is a significant characteristic of the true reflex, one that distinguished it from both a twitch and a voluntary action. With the knee jerk, however, the stimulus and response times characteristically appear to overlap. In a leg held passively with its joint at a right angle, percussion of the patellar tendon evokes an immediate, explosive extension of the lower limb, without apparent delay. Closer examination reveals an almost

insignificant latent period measuring only 10 sigma (1 sigma = 1/1000 second), whereas the shortest latent periods among indubitable reflexes are usually many times this value. Reflex blinking, for example, has a latency of 45 sigma, and among limb reflexes, the shortest latent period known was 17 sigma. In fact, as Sherrington reported, "the latency for the knee jerk is but little longer than that for direct stimulation of the extensor (muscle) itself." (IANS, 87) For many leading physiologists, most notably Westphal and Waller, the very short latent period of the knee jerk was sufficient to exclude it from being considered a true reflex. Its many names bear testimony to its once troublesome status: the "knee-phenomenon", the "knee twitch", and the name still in use today, clinically, the "knee jerk".

The problem was both theoretical and empirical. No overall theory was at hand to explain what goes on in a reflex, to explain latency, for example, and there was a general absence of familiarity with the observable features of indubitable reflexes. Before Sherrington, no one really knew how a reflex worked, and no systematic study of reflexes had been attempted. Sherrington's systematic approach soon revealed that the extensor-thrust reflex had the same latent period as the knee jerk -- a fact previously overlooked. But of greater importance, his theory of the integrative action of the nervous system advanced the physiological principles that *explain* reflex behaviour, as well as its apparent anomalies. With the theory of integrative action, Sherrington was able to explain the anomaly of muscle tonus as being a true reflex, and he coupled these results with the theory to explain why the anomalous knee jerk is also a true reflex, which he termed the "patellar reflex".

The poverty of pre-Sherrington physiology is akin to that of current day psychology's about the 'conditioned reflex', discovered by Ivan Pavlov (1849-1936). Without an overall theory explaining precisely what occurs in a Pavlovian reflex, it cannot be determined whether in fact it is a reflex, and thus we are faced with the same kind of theoretical trouble which plagued pre-Sherrington physiologists about the patellar-reflex. Allow me to elaborate.

Stedman's medical dictionary defines reflex as a "response"; "an involuntary movement or exercise of function in a part, excited in response to a stimulus applied to the periphery and transmitted to the nervous centers in the brain or spinal cord." By itself, the definition is not sufficient, but it does indicate a key concept, namely, that a reflex is an *involuntary* response. The definition provides for a distinction between reflexes that depend upon nervous centers located in the brain, and those that do not, those with nerve centers in the spinal cord. Thus it leaves open the possibility of a conditioned reflex, since these phenomena evidently depend upon cephalic centers.

Conditioned responses are learned. Yet, as it is held, once they are learned, they cannot be controlled or prevented from occurring by the subject who is conditioned. For this reason, the conditioned reflex is considered involuntary, and therefore truly reflex. To establish a conditioned reflex an animal, typically, is trained to associate one variable, foreign and unnecessary to its environment, with the occurrence of a second, familiar variable, preferably one that is also essential to its well-being. Food is the usual second variable, and virtually anything can serve as the

first; Pavlov used a ringing bell. Training proceeds with the introduction of the second variable, which itself must be sufficient to evoke the desired response. Training terminates, and a conditioned reflex is established, once the desired response can be evoked by the first variable alone, the foreign variable. Pavlov conditioned his dogs to salivate and secrete gastric juice, apparently without control, upon the ringing of a bell, without the introduction of food.

Thus far, my factual sketch of Pavlov's experiment is straightforward and unproblematic. The epistemological trouble arises, however, once an interpretation of the results is attempted, once we attempt to apply theory to fact. Are we to conclude that the dogs responded as if the bell were food, or did they respond in *anticipation* that food would be forthcoming? The first interpretation is easily ruled out. If Pavlov's dogs mistook the ringing of the bell to be food, one should expect to find a Pavlovian dog satisfied without the introduction of real food, and even refuse real food. But neither happens. Common sense appears to favour the second interpretation, yet the cost is considerable. If dogs can anticipate, they don't respond simply. The simplicity of the stimulus-response structure of an *involuntary* reflex is mediated by anticipation, and thus the black box of dog psychology. For example we cannot be certain that the dogs do not salivate to get the food, rather than because of the food; in other cases, dogs evidence goal-directed behaviour, such as oestrus behaviour, hunting, establishing territory, and fighting. The problem is one of being unable to decide whether the response is an involuntary effect, or a voluntary cause. We can't decide, because we cannot trace the course of events that occur in the conditioned 'reflex' immediately after the application of the

stimulus; the exact events between the stimulus time and the response time are hidden. We are without a theory that explains what these events are, and why they occur necessarily, rather than not. The physiological and psychological status of Pavlov's discovery is therefore uncertain and unknown. The same problem faced pre-Sherrington spinal physiologists, who were without a theory to explain the nature and causal necessity of spinal reflex events. Psychology awaits the arrival of its Sherrington.

To uncover the hidden spinal reflex, Sherrington began an exhaustive study of vertebrate neuromuscular anatomy, a study lasting ten years. He mapped nerve pathways from receptive surfaces to muscles in the rhesus monkey, the dog, the cat and the frog. The completion of his anatomical studies marked the end of what he considered a tedious, but necessary, beginning. To uncover the actual mechanism of the spinal reflex, Sherrington had to devise methods to measure the contribution made by each part in the process. Although a tap on the patellar tendon will evoke the patellar reflex, an accurate measurement is not possible unless the response can be *quantified* and measured against a specific quantity of stimulus. To obtain accurate results, in the case of the patellar reflex, Sherrington used weights instead of taps to pull the patellar tendon, measured the electrical excitation of each neurone in the reflex with a voltmeter, and used the *myograph* he invented to record muscle contraction; the myograph consists of an inked needle that marks a rotating drum according to the stress placed upon it by the muscle to which its opposite end is attached.

Because the nerve centers of spinal reflexes are located in the spinal cord, Sherrington was able to negate any influence the

brain has on a spinal reflex by removing the cerebrum (decerebrate animal), or the entire head (decapitate or spinal animal). By doing either, he eliminated the ubiquitous stimuli of the animal's normal environment, and was thereby able to effect only those stimuli necessary to a given experiement. Under these extreme conditions, it is still possible for an animal to stand, walk or gallop upon a treadmill, scratch itself, withdraw an injured limb, squat to defecate, and respond to pulls on its patellar tendon. Each of these activities is entirely reflex, entirely spinal.

Sherrington's physiological investigation was limited largely to the study of lower vertebrates (e.g. dogs and cats), because as one ascends the animal hierarchy, reflex activity becomes more and more dependent upon the influence of centers located in the brain. Yet in studying the function of spinal reflexes in lower vertebrates, Sherrington believed he was investigating neuromuscular and neurological elements that are basal to all vertebrates, at all levels of nervous activity. My attention is presently restricted to Sherrington's study of each component in the spinal reflex and to his discoveries of what contribution each component makes to motor coordination, or to the integrative action of a reflex.

3. MOTOR COORDINATION

Only three levels of coordination are the subject of the present chapter: first order coordination; second order coordination, and motor solidarity. The integrative action of the nervous system also produces a fourth and fifth level of complexification, respectively, psychical and psycho-physical integration, but my discussion of these

is reserved for the Appendix, where I consider whether Sherrington was a mind-body dualist.

What is a reflex? It was indicated earlier that a reflex is an *involuntary* reaction or response. Sherrington would not have been content with this definition, because it omits the most remarkable feature of reflex action, the feature Sherrington believed is the most difficult and important to explain, namely, motor *coordination*. A reflex is a complex, involuntary, coordinated response. Sherrington recognized two levels of motor coordination: *the degree necessary to render a response appropriate to a given stimulus*; for example, motor coordination must provide for the animal's ability to scratch when bitten by a flea, rather than squat; and the *degree of coordination necessary for the compounding of reflexes* which use the same limb for opposed purposes e.g. the extensor-thrust reflex versus the withdrawal reflex. Such sophisticated responses are remarkable, because they are not dependent upon centers higher than the spinal cord, i.e. precisely because the responses are coordinated, yet involuntary. For Sherrington, the definition of "reflex" should be carried a step further. A reflex is an involuntary, coordinated, *purposive* response. But leaving the property of purposiveness aside for the moment, how are both levels of motor coordination possible with the spinal reflex?

Sherrington's ten year anatomical study disclosed the general feature that a reflex-arc (see fig. 1) is composed of an inward path, a central organ, and an outward path. The inward path begins at a receptor organ located within a receptive field of such organs, and extends along a *sensory* (afferent) neurone leading to the central organ. The incoming path, composed of receptor organ and afferent

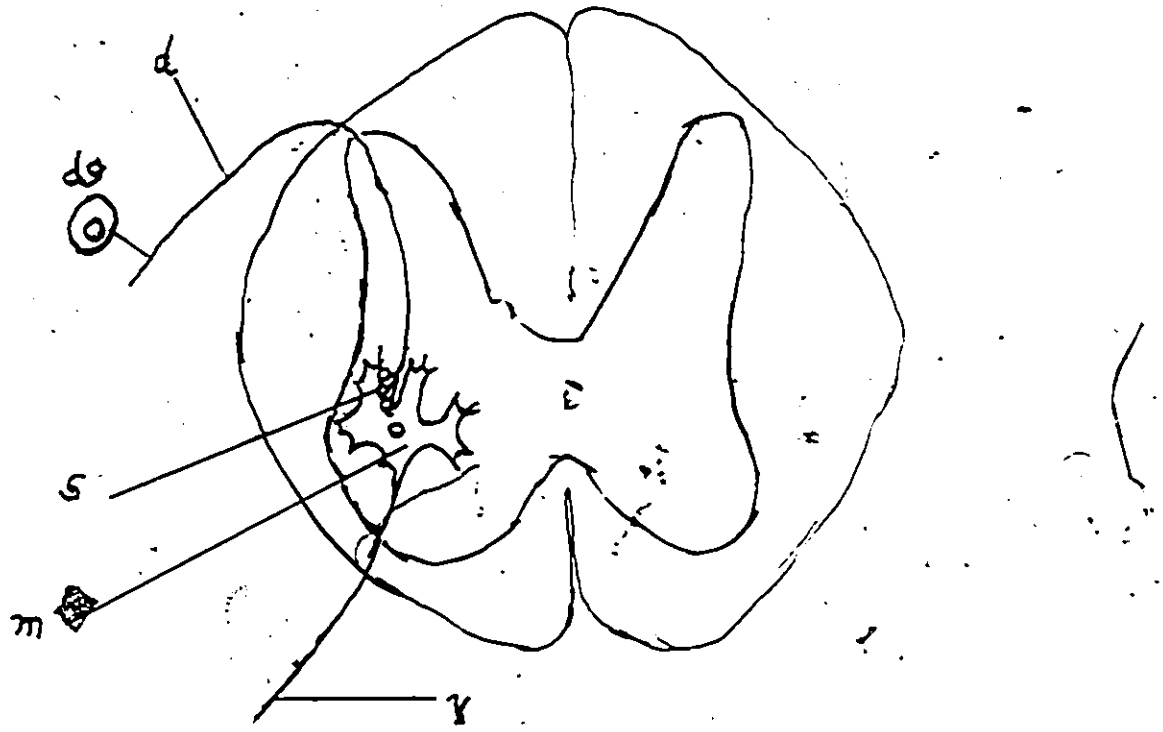


Figure 1

A diagram of a simple reflex-arc in the spinal cord.
d, dorsal nerve-root
v, ventral nerve-root
dg, ganglion of dorsal root
m, motor-nerve-cell in ventral horn
s, region of synapse

(from *Reflex activity of the Spinal Cord*, C.S. Sherrington, 1932, page 2.)

neurone, is what Sherrington calls, the "*afferent-arc*". The central organ of spinal reflexes is the grey matter of the spinal cord. The outgoing path begins at the central organ and proceeds along a *motor* (efferent) neurone that leads to an effector organ, either a gland or a muscle. The efferent neurone plus the effector organ constitute what Sherrington termed, the "*efferent-arc*". Destruction or loss of any of these structures in the intact animal results in various degrees of reflex in coordination. Forthwith we can examine the contribution to motor coordination afforded by the receptor organ.

3.1 THE RECEPTOR ORGAN

Between the booming, buzzing confusion of the external world, and the orderly composition of an animal, stand its receptor organs. The receptor organ has two functions which contribute to coordination: selective excitability, and conduction. A receptor is selectively excitable in respect to both the *quality* and the *quantity* of the stimulus impinging upon it. First, a receptor can respond to a stimulus only if the stimulus is of a specific kind. Thus, some receptors respond only to changes in temperature, i.e. heat receptors; others are capable of responding only to noxious stimuli, e.g. pain receptors; and others respond only if stimulated mechanically, e.g. stretch receptors. Second, if a stimulus, whatever its kind, is not of a sufficient strength, no impulse is imparted to the afferent neurone, the next link in the chain. If a stimulus is both *appropriate* and *adequate* it causes the receptor to conduct an impulse to the afferent neurone. A receptor organ has the capacity to discharge either a single impulse, or a train of impulses. Stimulus strength alters the frequency of the discharges, and secondarily, a high frequency alters

the size of each discharge. Of its two functions, selective excitability contributes more to coordination in the reflex; the receptor 'decides' *which* reaction takes place. Once the receptor discharges an impulse, the afferent neurone is largely responsible for *what* characteristics the response shall have.

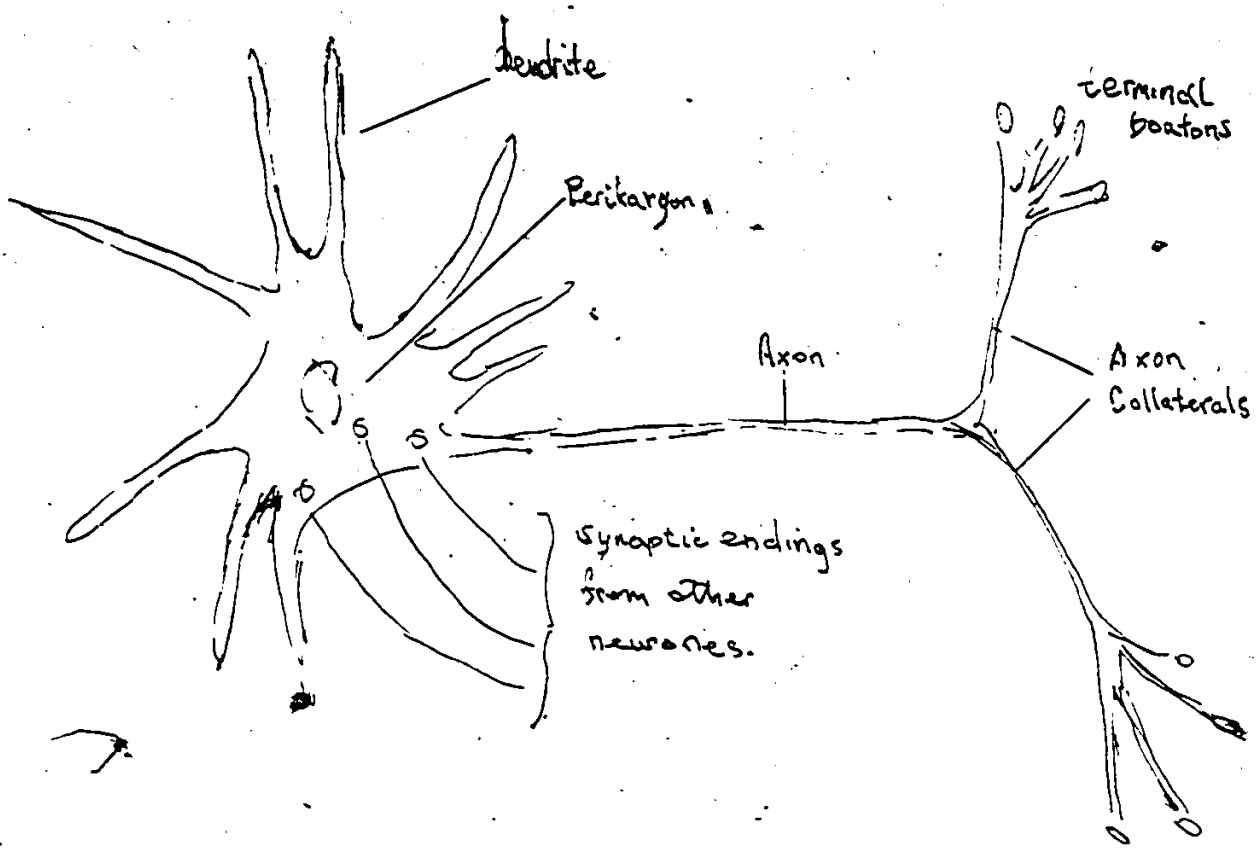
An example would be useful here to demonstrate the contribution selective excitability makes to coordination in the reflex. Consider the case of the powerful extensor-thrust reflex versus the withdrawal reflex. Pressure applied by the investigator's thumb to the paw cavity between the foot pads of a dog's hind limb evokes the extensor-thrust reflex. Such pressure is similar in kind to the pressure the cavity experiences when the hind paws are required to bear the weight of the animal as it runs or prepares to leap forward. Once the pressure has reached a certain threshold, the reflex thrust of the hind limbs springs the mass of the animal forward. If, however, a different kind of stimulus is applied to the same area of the foot, such as that which results from the application of a needle point, the withdrawal reflex is evoked, and the hind limb, rather than thrusting against the stimulus, draws away from it, the limb remaining suspended against the abdomen until the stimulus is removed. The receptor's contribution to coordination is obvious: disaster would ensue if the pressure receptors were selective in favour of pain as well as pressure, for then either kind of stimulus could evoke the extensor-thrust reflex. Without the receptor's capacity for selective excitability, responses would not be appropriate to stimuli, and an animal would be greatly endangered by otherwise manageable changes in its environment.

During Sherrington's inquiry into reflex coordination, uncovering the microstructures of a reflex was not usually a straightforward matter of dissection and observation, simply because the structures are too small to observe and handle with conventional microscopes and equipment. Yet the picture Sherrington gave us of the neurone, the synapse, and the neural arrangement of afferents and efferents was sufficient to account for the behavioural processes he studied. Ingeniously, Sherrington used physiology as a main tool for determining the microstructures underlying reflex action. Reasoning from effect to cause, he inferred the anatomical arrangements that must be present to bring about the coordinative action of the reflex. I will later describe this method of reasoning as one use of functional explanation.

3.2 NEURONE/SYNAPSE

Certain structural and functional features are common to both afferent and efferent arcs e.g. the neurone/synapse. In the opening pages of his *Integrative Action of the Nervous System*, 1906, Sherrington remarks, "The unit reaction in nervous integration is the reflex, because every reflex is an integrative reaction and no nervous action short of a reflex is a complete act of integration." (7) If the reflex is the nervous system's physiological unit, the neurone is its anatomical unit.

A neurone consists of a cell-body (or perikaryon), dendrites, and an axon. (See fig. 2) Dendrites increase the surface area of the perikaryon, upon which impinge stimuli from physiologically antecedent neurones. An axon is a long finger-like projection of the perikaryon through which the cell's own impulses funnel and travel by the axon



A neurone; explanation in text, p. 70 - 71. Figure 2

to impinge on the surfaces of subsequent neurones. Except for afferent neurones whose cell bodies are located peripherally, the nerve cells comprising the various reflex-arcs are grouped together centrally, to form the grey matter of the spinal cord. In the grey matter, impulses from one neurone impinge on the next neurone in the reflex-arc through terminal boutons or rings of microscopic size. The grey matter is the site of the all important *synapses*.

Sherrington introduced the term "synapse" to denote the functional membrane-to-membrane contact of a nerve cell with another nerve cell, an effector cell, or a sensory receptor cell. The width of the gap across a synapse is usually between 150 to 500 angstroms.

The simplest spinal reflex has one synapse between its incoming sensory neurone, and its outgoing motor neurone. The majority of spinal reflexes consist of an afferent neurone, an internuncial neurone, and an efferent. Conduction in the reflex-arc is, contrary to nerve trunk conduction, *intercellular* as well as *intracellular*. Significant aspects of coordination are the direct result of intercellular conduction at the synapse, conduction by contact, not by continuity.

Intercellular conduction was a novel concept that did not gain unanimous acceptance among physiologists until the second decade of our own century. Sherrington's reasons for asserting that conduction in the reflex is intercellular deserve consideration before we examine its significance to coordination. By far, Sherrington's most vocal opponents were the so-called Reticular theorists, who held that nerve impulses can only be conducted through *unbroken* nerve channels. It is easy to appreciate the difficulty Reticularists had with Sherrington's novel theory. How is conduction possible in a reflex-arc, given these synaptic 'breaks' thoughout? After all, in every known circuit in the electrical world, conduction ends precisely when the circuit breaks. Moreover, how can a system that is fragmented into a billion neurones provide for functional unity? How is functional unity possible without a continuous nerve network, i.e. without anatomical unity?

Certainly Sherrington's greatest problem in this regard was the absence of unequivocal histological evidence to confirm the precise structure of the synapse. He writes: "The minute structure of these points of contact between the nerve cells, 'synapses', is by no means easy to determine, because shrinkage and distortion of the

very delicate elements involved must necessarily accompany fixation, imbedding, and staining." (RASC, pp. 13-14) He was forced to seek physiology as his source of anatomical knowledge, a point I will develop in Chapter IV on the role of functional explanation in science.

Sherrington compared conduction along a reflex-arc with conduction in the long myelinated nerve trunks of the voluntary nervous system, and discovered a striking qualitative difference, which he elevated to the status of a law. This was the first clue about the existence of the synapse, and therefore also of neurones. Nerve-trunk conduction is *antidromic*; reflex-arc conduction is dromic, meaning its impulses are conducted in one direction only: cellulipetal in dendrites and cellulifugal in axons. This principle characteristic of conduction in the reflex is called the *law of forward direction*. Only the synapse, with irreciprocally permeable surfaces of separation, could explain the chief physiological property of reflex-arc conduction. Sherrington argued thus:

> "Nerve-fibres are known to conduct impulses equally well in either direction, whereas no reflex movement can be evoked by stimulation of the central stump of a cut ventral spinal root. This 'law of forward direction' cannot well be attributed to the cell-bodies contained in the grey matter, since dorsal spinal root ganglia conduct in both directions. Nor ... is it the property of those cell-processes which are met in the peripheral nerve. The junctions between successive neurones are the only other structures interpolated in the reflex-arc to which we can ascribe such blockage of 'antidromic' conduction." (IANS, 13)

As well as accounting for the qualitative difference between reflex-arc conduction and nerve trunk conduction, only the synapse could explain an observed spectrum of quantitative differences between the two types of conduction. At the outset of his

Integrative Action, Sherrington listed the following qualitative differences between the two, and devoted most of his "Lectures" to their elaboration and explanation. He writes:

> "Conduction in reflex-arcs exhibits: (1) slower speed as measured by the latent period between application of stimulus and appearance of end-effect, this difference being greater for weak stimuli than for strong; (2) less close correspondence between the moment of cessation of stimulus and the moment of cessation of end-effect, i.e. there is a marked 'after-discharge'; (3) less close correspondence between rhythm of stimulus and rhythm of end-effect; (4) less close correspondence between grading of intensity of the stimulus and the grading of intensity of the end-effect; (5) considerable resistence to passage of a single nerve-impulse, but a resistence easily forced by a succession of impulses (temporal summation); (6) irreversibility of direction instead of reversibility as in nerve-trunks; (as noted) (7) fatigability in contrast with the comparative unfatigability of nerve-trunks; (8) much greater variability of the threshold value of stimulus than in nerve-trunks; refractory period, "bahnung", inhibition, and shock, in degrees unknown for nerve-trunks; (10) much greater dependence on blood-circulation, oxygen (Verworn, Winterstein, v. Baeyer, etc.); (11) much greater susceptibility to various drugs -- anaesthetics." (IANS, pp. 13-14)

On what causes these differences, Sherrington proposed:

> "If the conductive element of the neurone be fluid, and if at the nexus between neurone and neurone there does not exist actual confluence of the conductive part of the one cell with the conductive part of the other, e.g. if there is not actual continuity of physical phase between them, there must be a surface of separation. Even should such a membrane visible to the microscope not appear, the mere fact of non-confluence of one with the other implies the existence of a surface of separation. Such a surface might restrain diffusion, bank up osmotic pressure, restrict the movement of ions, accumulate electric charges, support a doubled electric layer, alter in shape and surface tension with changes in difference in potential, alter in difference in potential with changes in surface tension or in shape, or intervene as a member between dilute solutions of electrolytes of different concentration or colloidal suspensions with different signs of charge. It would be a

> mechanism where nervous conduction, especially if physical in nature, might have grafted upon it characters just as those differentiating reflex-arc conduction from nerve-trunk conduction. For instance, change from reversibility of direction of conduction to irreversibility might be referable to the membrane possessing irreciprocal permeability ... It seems therefore likely that the nexus between neurone and neurone in the reflex-arc, at least in the spinal arc of the vertebrate, involves a surface of separation between neurone and neurone; and this as a transverse membrane across the conductor must be an important element in *intercellular* conduction. The characters distinguishing reflex-arc conduction from nerve-trunk conduction may therefore be largely due to intercellular barriers, delicate transverse membranes in the former." (IANS, pp. 16-17)

Sherrington's physiological studies thus bound him to the belief in synapses and neurones, a belief that later proved to be correct.

Sherrington presented the Reticularists with a markedly different model of reflex conduction based on his law of forward direction, as well as another equally important law of neuronal conduction in the reflex. A nerve fiber is the executive process of a neurone that exists together with similar axons of other neurones to form a nerve. Each fiber conducts electrical impulses independently of its fellow fibers in the nerve. Stimuli are either *subliminal*, and evoke no response in any nerve fiber; *liminal*, and excite a few; or *maximal*, and excite the entire nerve. In contrast, the response of any neurone fiber is either maximal or nil; which is Sherrington's second major principle of conduction in the reflex, the *all or nothing principle*.

This principle suggests the analogy that a neurone is akin to a trail of gun powder. Both the neurone and the trail are

forms of stored energy. If either is caused to release its own energy, it is an all-or-nothing situation. Furthermore, the amount of energy released by both is considerable when compared to the amount imparted to either by an appropriate cause. A small flame is sufficient to cause a spectacular release of energy in the gunpowder, and similarily, a tiny flea placed in the ear of a dog will evoke a violent shaking of the dog's head which persists until the flea is ejected.

If a neurone responds to an adequate stimulus by releasing its own energy, one should expect that its capacity to respond to a volley of such stimuli is limited by the rate at which it can replenish its energy stores. This is precisely what Sherrington found. After a neurone discharges an impulse, which discharge requires only 0.9 sigma, there occurs in the neurone an *absolute refractory period*, lasting a full sigma, during which no stimulus, however great, can evoke a response; and for a further 5 sigma a *relative refractory period* is evident, during which a very powerful stimulus can evoke only a weak response. Evidently, during the absolute refractory period the neurone's energy stores are depleted, and during the relative refractory period they are being replenished.

Successive nerve impulses, which create a tetanic muscular contraction characteristic of normal muscular responses, cannot "fuse into a steady flow, like the conduction of water in a pipe or an electric current in a wire, but may be compared to the repetitive discharge of bullets from a machine-gun." (RASC, 4) Sherrington's machine-gun analogy embodies both the concept that neuronal discharges are discrete packets of energy (all-or-nothing principle), and the

notion that the discharges are unidirectional (law of forward direction).

Having outlined the chief features universal to the neurone and to conduction in the reflex, discussion can now turn to the structural and functional properties unique to an afferent neurone to show how these neurones contribute to coordination in the reflex.

3.3 AFFERENT NEURONE

Adequate stimuli from a peripheral receptor impinge upon the dendrites of an afferent neurone, which relays an electrical impulse down its axon that enters the grey matter dorsally. After entering the cord, the axon bifurcates headwards and tailwards with each section divided again many times to form collateral branches, whose endings are *boutons terminaux*. (See fig. 2) The course of a fibre within the grey matter probably never exceeds a spinal segment (i.e. the width of a single vertebra), and no afferent neurone trespasses across the invisible midline; a cross-reflex (one whose stimulus point is on the opposite side of the midline to its responding limb) therefore always involves an internuncial neurone. Special features of an afferent neurone occur at its peripheral end and its central end.

An adequate stimulus effects an immediate explosive discharge of stored energy in the afferent neurone which, in turn, stimulates the next neurone in the chain. Although subliminal stimuli evoke no such response, they do effect in the afferent neurone what Sherrington calls a "*local excitatory state*" (l.e.s.). The l.e.s. is a state of subliminal excitation that remains at the

neurone's peripheral end for 1 sigma. If a second subliminal stimulus impinges upon the neurone within that period, the two are summed to create a greater state of local excitation. If, as a result of their *summation*, the l.e.s. exceeds the "threshold", the neurone discharges an impulse. The l.e.s. differs from a regular nerve-impulse in four respects:

1. It remains localized at the stimulated point, and does not conduct.
2. It does not obey the all or nothing principle; rather, the quantity of the l.e.s. is a product of the quantity(ies) of the subliminal stimulus(i).
3. It causes no refractory period in the neurone, which provides for the fourth difference.
4. It is capable of summation.

The l.e.s. and the nerve-impulse are the only excitatory conditions of the afferent neurone.

At the terminal or central end of an afferent neurone are its collateral branches which innervate the grey matter of the spinal cord. The boutons terminaux of these collaterals either *excite* or *inhibit* the next neurone, e.g. an efferent neurone. Regarding the afferent boutons Sherrington writes: "So far as is known ... no clear morphological difference distinguishes the 'boutons' into two classes, although ... central excitation and central inhibition are two distinct processes occurring in the grey matter." (RASC, 14) Again we see that Sherrington is faced with limitations imposed by the technology of his time; but once again he used physiology to circumvent the problem: "The striking correspondence observed between the reflex inhibition and the reflex contraction, when examined in one and the same reflex-type, allows the inference that

nerve-fibres from the receptive field of the reflex each divide in the spinal cord into end-branches (e.g. collaterals), one set of which, when the nerve-fibre is active, produces excitation, while another set, when the nerve-fibre is active, produces inhibition. The single afferent nerve-fibre would therefore in regard to one set of its terminal branches be *specifically excitatory*, and in regard to another set of its central endings be *specifically inhibitory*. It would in this respect be duplexed centrally." (IANS, pp. 103-04)

Since it is the efferent neurones that exhibit the states of *central excitation* and *central inhibition* (imparted to them by the afferent neurones), a study of these effects is appropriately reserved for the upcoming discussion, on the efferent neurone.

If the receptor organ 'decides' whether a stimulus evokes a response, the afferent neurone determines what form the response will take. That is to say, the *rhythm of response* is largely a function of the afferent neurone. The rhythm of a reflex response is its characteristic movement or absence of movement. The rhythm may be *static* as is the case with postural reflexes, or *aphasic*, as in the case of the scratch reflex. In determining the rhythm, the afferent neurone contributes to coordination by ensuring that a reflex response is appropriate to its stimulus.

The explanation of the capacity to render appropriate responses can be stated simply. Afferent neurones show great diversity in respect to thresholds of excitability, frequencies of discharge, and refractory periods. Such variations between afferents give rise to their different response characteristics, which control the rhythm of response in a given reflex. This determinacy renders, for example, the afferents of the withdrawal reflex incapable of evoking a response

rhythm that would characterize the scratch reflex, and so on. The afferent neurone is thus the main determinant of first order coordination, the coordination within a single reflex which gives it appropriate response characteristics.

3.4 EFFERENT NEURONE

The perikaryon or cell body of an efferent neurone is located in the grey matter. The axon leaves the grey matter ventrally and extends into an effector organ, either a gland or a muscle. Contraction of skeletal muscles is the expression of reflex centers Sherrington considered, and thus muscles are the effector organs he studied. Before and after entering a muscle, the axon of an effector neurone splits into collateral branches whose boutons innervate, on the average, 100 muscle fibres. Muscle fibres or strands are arranged in parallel and, typically, the proximal end is attached to bone and the distal end to a tendon. An efferent neurone, together with the 100 muscle fibres which it innervates, is what Sherrington refers to as a *motor unit*. A typical muscle is a collection of tens of thousands of muscle fibres, and thus a muscle consists of hundreds of motor units.

Physiologically, a motor unit is a sub-division of a reflex-arc, the true unit of the nervous system because of its capacity for coordinative (integrative) action. The features of coordination attributable to the afferent arc (which includes the synapse) can be determined by comparing a *motor response* with a *reflex response*. Stimulation of the central end of an efferent neurone (eg. the sciatic nerve) causes a motor response, whereas the same stimulus applied to

a relevant afferent neurone evokes a modified, reflex response. This modification is the product of features imposed on the response by the afferent-arc. In addition to disclosing features of the afferent-arc, such a comparison discloses features about the relation between afferents and efferents in a reflex-arc.

The synapse between an afferent and an efferent is most characteristically a meeting place of many afferents and many efferents. Sherrington found that the afferents outnumbered the efferents 5 to 1, and he termed this arrangement the *principle of convergence*. "This principle of convergence is the basis of co-ordinated function in the nervous system." (RASC, 25) Two physiological features of convergence, *fractionation* and *occlusion*, supported Sherrington's histological findings, and pointed the way to the contribution of convergence to coordination in the reflex.

Let us first explain *fractionation*. Stimulation of an afferent neurone evokes a reflex response characterized by various degrees of contraction in the muscles involved. The degree of contraction can be measured by attaching the muscle's distal end to a meter, graduated in terms of weight. Sherrington observed that beyond a certain point, any increase in stimulus strength failed to evoke further contraction in the muscle. This indicated that for any afferent there is a limit to its effect, called its "maximal reflex response". However, direct stimulation of the relevant efferent indicated that the muscle tension of the same muscle can exceed the above limit. Consider the results Sherrington obtained by examining one muscle and all of its afferent neurones:

M. Tibialis Anticus. Maximum motor tension 2160 grams.

Afferent N. stimulated	Tension of maximal reflex tetanus in grams.	Reflex tension expressed as percentage of maximal motor tetanus.
Internal saphenous	800	32
Superficial obturator	165	6.7
Deep obturator	400	16
N. to quads & sartorius	1190	44
Musculo-cutaneous branch of peroneal	1700	69
External plantar	1240	50
Internal plantar	1330	54
Small sciatic	680	28
Hamstring	565	23
N. of sural triceps	300	12
	TOTAL 8370	

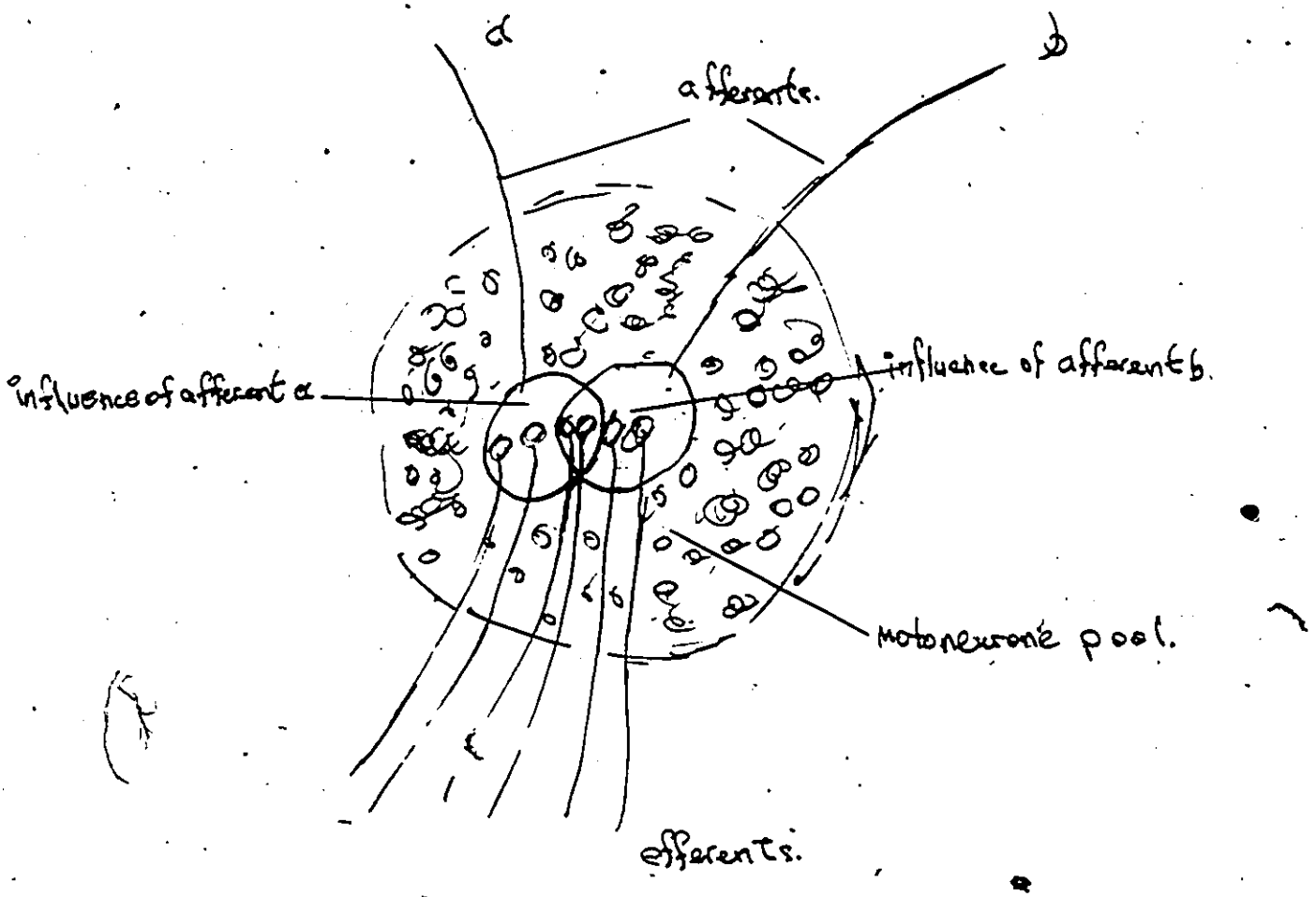
(from RASC, p. 24)

The table shows that no afferent is able to evoke more than 70% of the maximum motor response in the M. Tibialis Anticus; indeed the average is below 35%. This indicated to Sherrington that only a fraction of the total number of motor units comprising a muscle are accessible to any particular afferent neurone. He called this feature of the neural arrangement, "fractionation".

To understand *occlusion*, the second feature, compare the total tensions evokable by all the afferents, with the maximum motor tension (i.e. compare 8370 gr. with 2160 gr.). Since the total exceeds the maximum motor tension, concurrently stimulated afferents must block or occlude one another's effect on the efferent they share.

Fractionation and occlusion suggested to Sherrington the idea of a *motoneurone pool*, in which the impulses from many afferents

converge to impinge upon common and uncommon efferents. (See fig. 3)



A motoneurone pool; explanation in text, p. 82 - 83. Figure 3

(from RASC, p. 27)

Figure 3 shows that afferents a and b have access to only a fraction of the outgoing motoneurones. It also shows how if either afferent stimulates its efferents maximally, subsequent incoming volleys from the second afferent are able to excite maximally only 50% of that afferent's set of efferents, since the remaining 50% are being maximally stimulated by the first afferent. If b is the second stimulating afferent, b is said to suffer a 50% occlusion from

afferent a.

Thus we see how the proposal of fractionation and occlusion cohered with the principle of convergence of afferents upon efferents. We will make use of this line of reasoning in Chapter IV. It must be emphasized that under normal conditions, a single afferent does not always discharge impulses in succession which summate to cause excitation in an efferent to rise to a threshold level. What is more typical is for afferent impulses from a number of afferent sources to summate their effects on the motoneurone. Afferents do not always occlude one another, since occlusion occurs only if the shared motoneurone has reached its threshold of excitation. Under extreme conditions, fractionation and occlusion are more evident than under normal conditions, yet, of course, convergence, which these phenomena indicate, is characteristic of the arrangement of the nervous system under both conditions. The excitatory states imparted to the motoneurone by its afferents, in the normal interplay of reflexes, is appreciated best with the understanding that a motoneurone is a *final common path* for a number of afferents.

The principle of convergence is companion to the *principle of the final common path*. Afferents, *private* paths, converge upon internuncials, *common* paths, which converge upon motoneurones, *final common* paths. The principles of convergence and the final common path (FCP) depict the same concept of neural arrangement, but the former emphasizes the relation of afferents to efferents, whereas the latter emphasizes the relation in reverse. Sherrington describes the relation in the following passage:

> "Reflex-arcs show ... the general feature that the initial neurone of each is a *private* path exclusively

> belonging to a single receptive point (or small group of points); and that finally the arcs embouch into a path leading to an effector organ; and that their final path is common to all receptive points wheresoever they may lie in the body, so long as they have connexion with the effector organ in question. Before finally converging upon the motoneurone the arcs converge to some degree. Their private paths embouch upon *internuncial* paths common in various degrees to groups of private paths. The terminal path may, to distinguish it from internuncial common paths, be called the *final common path*. The motor-nerve to a muscle is a collection of final common paths." (IANS, 118).

Sherrington draws two consequences from the above arrangement for reflex coordination. The first supports my earlier remarks on the contribution of the afferent neurone.

> "Certain consequences result from this arrangement. One of these seems to be the preclusion of essential qualitative difference between nerve-impulses arising from different afferent nerves. If two conductors have a tract in common, there can hardly be essential qualitative difference between their modes of conduction; and the final common paths must be capable of responding with different rhythms which different conductors impress upon it. It must be to a certain degree aperiodic. If its discharge be a rhythmic process, as from many considerations it appears to be, the frequency of its own rhythm must be capable of being at least as high as that of the highest frequency of any of the afferent-arcs that play upon it; and it must be able also to reproduce the characters of the slowest." (IANS, 118)

Thus we see that the responsibility for determining a reflex's rhythm of response must be a function of its afferent neurone, since the FCP or efferent motoneurone must have the capacity to respond to the rhythmic demands of many reflex, afferent-arcs.

The second consequence Sherrington draws sheds light on the contribution of the efferent-arc.

> "A second consequence is that each receptor being dependent for final communication with its effector organ upon a path not exclusively its own but common to it with certain other receptors, such nexus

> necessitates successive and not simultaneous use of the common path by various receptors using it to *different or opposed* effect. When two receptors are simultaneously stimulated, each of the receptors tending to evoke reflex action that for its end-effect employs the same final common path but employs it in a different way from the other, one reflex appears without the other. The result is *this* reflex or *that* reflex, but not the two together." (IANS, pp. 118-19)

That is to say, the FCP must provide for its compound employment by many arcs, some of them *allied*, and others of them *antagonistic*. How this second level of coordination is achieved, how the compounding of reflexes is possible, is a function of the excitatory and inhibitory states imparted to the FCP by allied and antagonistic afferent neurones.

An incoming afferent impulse impinges upon the central ends of a number of efferent neurones causing them either to be excited or inhibited, depending on whether they are affected by excitatory or inhibitory boutons. Occasionally, the incoming impulse of an excitatory bouton is sufficient to cause the efferent to discharge an impulse. More often however, the excitatory bouton causes what Sherrington calls a *central excitatory state* (c.e.s.) in the efferent. The c.e.s. is an enduring excitatory state of the motoneurone lasting 10 to 20 sigma after the subliminal stimulus of the afferent has ceased. The c.e.s. is akin to the l.e.s. of an afferent insofar as the c.e.s. is merely also a predisposition to discharge, and is not an actual discharge of an impulse. Moreover, like the l.e.s., the c.e.s. can be summated to threshold levels by subsequent (afferent) volleys. But in contrast the c.e.s. lasts 10 to 20 times the period of the l.e.s. Furthermore, unlike the l.e.s. the predisposition to discharge of the c.e.s. is not only affected by the

arrival time of subsequent volleys, but is also influenced by what Sherrington termed the central inhibitory state (c.i.s.).

Unlike both the l.e.s. and the c.e.s., the c.i.s. cannot cause a (moto)neurone to discharge an impulse. Rather, the c.i.s. inactivates the c.e.s. The c.i.s. outlasts the c.e.s. 5 to 10 times. On the other hand, like both the l.e.s. and the c.e.s., the c.i.s. can summate in a (moto)neurone to reduce its capacity to discharge an impulse even further. Although the exact nature of the inhibitory state was unknown to Sherrington, he thought that it cancelled an FCP's disposition to discharge an impulse by stabilizing the surface of its synaptic membrane.

Allied afferents reinforce each other's effect of an FCP, such that a volley from each alone may not be sufficient to evoke a reflex response, whereas their summation may achieve the excitatory threshold to cause a reflex response. Consider, for example, the left scratch reflex. Its receptive field is a saddle-shaped area of the left dorsal skin, whose perimeter follows the dorsal midline, from the back of the neck to the base of the spine, down the left flank, around the rib-cage, and back up to the base of the neck. Afferent-arcs of the left scratch reflex receive stimuli from this saddle-shaped receptive field. The degree of alliance between these afferent-arcs is a function of the distance between their receptive points, i.e. the closer the receptive points, the greater the alliance between the arcs. With regard to evoking a reflex response, normally the activity of a flea, for instance, is sub-threshold provided the flea remains in one place on the field and doesn't bite too frequently (i.e. with a frequency whose period is greater than the duration of

the c.e.s.). If the flea begins to crawl, however, it endangers evoking a threshold state of excitation in the FCP, because its crawl will excite more afferent-arcs which ally their states of excitation by contributing to the building up of the c.e.s. to a threshold level, that causes the FCP to discharge. The characteristic of alliance in the scratch reflex has been so persistent throughout evolution, that fleas are adapted to hop, not crawl, which greatly reduces the likelihood of their evoking the scratch reflex, and so risking destruction..

Antagonistic afferents inhibit one another by creating a c.i.s. in the FCP's they do not share. The unshared motoneurones impinged upon by each afferent in the motoneurone pool determine which of the competing afferents evokes a reflex response, since the shared FCP's are more likely to be in a threshold state of excitation. Which afferent-arc wins the use of the limb depends upon the number of efferents it can excite, and to what degree, over and above the number that are inhibited, and to what degree, by the antagonistic arc; and upon the number of efferents, employed by the antagonist, that it can inhibit, and to what degree. If two antagonistic reflexes are competing for the use of the same limb, the ensuing response may have characteristics of the recessive reflex, because it exerts an influence on shared motoneurones, but these characteristics will not be sufficient to render the dominant reflex incoordinate, and unrecognizable.

Because allied reflexes can cooperate their effects upon a group of FCP's, convergence to the final common path renders first order coordination (i.e. an appropriate response) possible; and

since antagonistic reflexes are forced to compete for the use of fewer FCP's, convergence to the final common path renders second order coordination necessary. Convergence to the final common path engenders the solidarity of an animal in regard to its reflex actions.

3.5 RECIPROCAL INNERVATION

The c.i.s. is decisive in determining which of two, or more, antagonistic reflexes gains temporary domination over an otherwise shared limb. For this reason, the c.i.s. is a crucial factor in the physiology of second order coordination. Without the c.i.s., antagonistic reflexes could be thrown into action simultaneously, to the demise of the animal. As I indicated, the histological source of the c.i.s. is the inhibitory bouton; the inhibitory boutons of antagonistic arcs synapse with the unshared FCP's of their opponents, to create in them a c.i.s. What remains for discussion is the role of the inhibitory bouton, and the c.i.s., in first order coordination.

A dominant reflex is required not only to inhibit the motoneurones used exclusively by its antagonist and to excite its own set of motoneurons, in addition, it must have the capacity to excite certain muscle groups of the limb it deploys while inhibiting others. Without this capacity, all muscles of a limb could be thrown into equal states of excitation, or inbibition, simultaneously; both effects are equally dysfunctional, and incoordinate. To prevent such a fate, the afferent boutons of a reflex-arc synapse *reciprocally* with efferent neurones innervating *flexor* muscles, that adduct the limb toward the midline, and with efferent neurones innervating *extensor* muscles, that abduct the limb away from the midline. The flexors are thus

inhibited when the extensors are excited, and vice versa. Such an arrangement of afferent boutons characterizes the simple flexor reflex, and the withdrawal reflex. Moreover, if the reflex is, like the scratch reflex, phasic, afferent boutons of both types synapse with each motoneurone in both flexors and extensors; the afferent neurones in such a reflex must be able to excite and alternately inhibit both the extensors and the flexors. Both types of neural arrangements, one that provides for *simultaneous* inhibition and excitation of flexors and extensors respectively and the other that provides for *successive* inhibition and excitation are examples of, what Sherrington termed, the *Principle of Reciprocal Innervation*. Reciprocal innervation is evident at yet a third level: extensor and flexor muscles are reciprocally innervated with respect to each other.

Reflexes do not exercise their effects on flaccid limbs; on the contrary, muscles resist usage. Under normal conditions, all extensor and flexor muscles evidence a basal state of contraction known as *muscle tonus*, the anomaly that long baffled physiologists before Sherrington. Muscle tonus is a reflex whose receptor organs, the so-called "stretch receptors", are located in the muscles themselves. Stretch receptors transmit signals to the cord which synapse with the FCP's of the same muscle to cause basal excitation, and the FCP's of the opponent muscle to cause basal inhibition. Flexors and extensors are therefore reciprocally innervated with respect to each other; each muscle exerts control over its own excitation and over its opponent's inhibition. For example, if an extensor is caused to stretch due to the contraction of its opponent flexor, the extensor

stretch receptors are caused to lengthen which increases their frequency of nervous discharge, and this in turn excites reflex contraction in the extensors and inhibits further contraction in the flexors. By means of reciprocal innervation, flexor and extensor muscles are self-regulatory. When not influenced by 'additional' reflexes, the flexors and extensors of a limb exhibit roughly equal degrees of contraction, which defines the limb's normal posture. Consequently, for example, if the flexor reflex is to deploy a limb, it must augment the degree of contraction in the flexor muscles, and inhibit the resistant muscle tonus of the corresponding extensor muscles. Once any such 'additional' reflex ceases to influence the limb, the muscle tonus reflex returns the limb to its normal posture.

Two points deserve emphasis. 1. Reciprocal innervation is necessary to all reflex action. It is a condition of reflex interaction, both allied and antagonistic; a condition of reflex efficacy; and a necessary condition of the muscles themselves. 2. *Adjustment in the quantity of contraction* in a muscle -- the sole expression of spinal reflexes examined by Sherrington -- is possible because excitation and inhibition are "coordinate reciprocal features in one united response". (IANS, 98) That is to say, excitation and inhibition are quantitatively antagonistic: the degree of contraction in a muscle is an expression of the quantity of excitation minus the quantity of inhibition imparted to its motoneurones, by all reflex-arcs involved in its reflex contraction.

Sherrington faced a comparatively easy next step to explain the anomalous knee jerk, once he had disclosed muscle tonus as a reflex action. The knee jerk is a reflex response of the quadriceps

muscle fibres to their stretching caused by percussion of the patellar tendon. The latent period of the patellar reflex is as short as it is because the reflex is superimposed upon (i.e. allied with) the ongoing reflex of muscle tonus in the quadriceps. A tap on the patellar tendon directly augments the reflex muscle tonus in the quadriceps and immediately inhibits tonus in the hamstring muscles; the sudden absence of resistance from the flexor muscle on the contraction of the extensor causes the lower limb to jerk forward in the characteristic manner. Conversely, Sherrington demonstrated that the latent period of the patellar reflex increases as the tonus of the quadriceps is gradually inhibited. Extreme inhibition destroys the knee jerk completely.

3.6 SUMMARY

My discussion until now has concerned the general principles of *coordinated* reflex behaviour Sherrington observed and developed, and I attempted to show how he derived them. My discussion indicates that the concept of reflex coordination entails:

1. The Principle of Selective Excitability
2. The Law of Forward Direction
3. The All-or-Nothing Principle
4. The Principle of Convergence
5. The Principle of the Final Common Path
6. The Principle of Reciprocal Innervation/Inhibition

Each of these physiological principles denotes a necessary condition, or set of necessary conditions, for coordinated reflex behaviour, which coordination is expressed as a predictable adjustment in the

quantity of contraction of limb muscles. The following chart shows to which anatomical structures the above physiological properties are attributed:

Chart II

1st Order Coordination:	response appropriate to stimuli	
structure	determines	function
A) receptor organ		1. selective excitability 2. conduction to afferent neurone
B) neurone/synapse		3. unidirectional conduction 4. threshold conduction
C) afferent neurone		response rhythm
2nd Order Coordination:	compound use of effector by reflexes of different or similar type	
D) efferent neurone (in conjunction with the physiological properties of the afferent neurone of: 5. convergence to a FCP 6. reciprocal innervation/ inhibition)		successive or simultaneous use of effector organ by antagonistic or allied afferent-arcs, respectively

4. MOTOR SOLIDARITY

The spinal animal we have examined in Chapter II was described by Sherrington as a "cartesian puppet", a wheelwork animal devoid of mind: "thoughts, feelings, memories, percepts, conations, etc. of these no evidence is forthcoming or to be elicited. Yet the animal remains a motor mechanism which can be touched into action in certain ways so as to exhibit pieces of its behaviour." (IANS, xi) As we have seen, these pieces of behaviour are reflexes that respond appropriately to adequate stimuli imparted to a receptive field. Further, the pieces of behaviour act harmoniously with one another, either simultaneously through alliance, or successively through reciprocal

inhibition (antagonism). Although Sherrington conclusively demonstrated the spinal cord's remarkable ability to effect both these grades of coordination, he also found the motor mechanism to be limited in the degree to which it can exhibit such coordination.

For example, the hind limb of a spinal animal evidences a scratch reflex that does not pinpoint the exact spot of an electric "flea", but only paws the general locale. Further, the reflexes of a spinal animal evidence weaker alliance and antagonism; Sherrington found that local stimuli evoke merely local responses. In other words, an investigator 'working' a spinal animal can approximate, but not duplicate, the precision of limb movement, and reflex harmony achieved by an animal that has not been "surgically mutilated". Moreover, an intact animal not only has greater coordination, its reflex responses are not *inevitable*. This feature is most apparent when reflex sequences are considered. For example, in a spinal dog reflex evacuation of the bowel is inevitably, and immediately, followed by reflex engagement of all four limbs which scratch backwards as if to earth-over the faeces. This sequence of reflex events is not inevitable with an intact dog, as pedestrians well know, which tells that its reflexes are modified under cerebral direction and control. (IANS, p. 305) Overall, in his comparison of reflex responses in the intact animal with those of the spinal animal, Sherrington found that with the former reflex precision is gained, whereas reflex neutrality (neutrality between reflexes) and reflex inevitability are lost.

The results of Sherrington's comparison thus pointed to a third property universal to animal behaviour, namely, motor coordination is augmented and regulated to engender *motor solidarity*: the integrative action of an intact nervous system evidences a receptiveness

to factors affecting the entire animal, and a responsiveness that engages the animal as a unified whole. While motor coordination is the remarkable achievement of the spinal cord, Sherrington saw motor solidarity as the "crowning achievement" of the brain.

Sherrington remarked that the spinal man is more crippled than the spinal frog. By this he meant that as one ascends the phylogenic scale, motor coordination depends more and more upon centers higher than the spinal cord. Yet, even with humans, most of the burden for motor coordination rests with the spinal cord. To effect movement of a limb, for example, the brain signals an agonist, which signal causes *automatic* inhibition of the corresponding antagonist, because the spinal circuit is such that efferent-arcs reciprocally innervate opposing groups of muscles. Imagine a world where the spinal cord is not so structured, yet where human brains are essentially what they are in our own world. Imagine what it would be like to live in such a world. To walk, one is required to think, and to will: 'place left foot before right by reciprocally inhibiting, in an appropriate sequence, the extensors and flexors of the left leg, while swinging right arm out by reciprocally inhibiting, in an appropriate sequence, the extensors and flexors of the right arm, then repeat the process, *mutatis mutandis*, to effect a second step using the right leg and the left arm while leaning forward on the left leg and retaining balance by keeping the right arm swung back', and so on. In such a world, sprinting would tax the best minds, hockey players would be geniuses, and a pianist could produce, at best, an extremely stilted *andante*. Indeed it is difficult to conceive of human culture developing at all in such a world, where physiological conditions

constrain the higher neurological, and therefore the psychological, centers to preoccupation with motor coordination. The point is that, in such a world, individuals are reduced to participants in activities that cannot simultaneously be assessed, adjusted and improved upon. In our own world, by contrast, the brain is provided with physiological conditions that free it from the responsibility for motor coordination, and which therefore enable it to augment and regulate motor coordination. The integrative action of the spinal system to effect motor coordination is thus a necessary condition of motor solidarity.

For the most part, lectures VII to X of the *Integrative Action* are concerned to describe the motor responses elicitable from an intact animal, and to explain the factors that produce the behavioural property of motor solidarity. Of particular concern is the contribution made by the brain. Sherrington's experiments on the brain copied the approach he had perfected in his original experiments to determine how the integrative action of the spinal system produces the two grades of coordination. Here as well, responses were elicited by applying an electric current directly to various areas on the surface of the brain; the brain's "functional topography" was thus determined. And second, by comparing such intact responses with the responses elicitable after various parts of the brain had been surgically removed, Sherrington determined the contribution each part makes to motor solidarity. Of course, owing to the complexity of the brain, Sherrington's account of motor solidarity could not approach the completeness of his account of motor coordination. Nonetheless, he was able to show which structures are involved in producing the

two general expressions of motor solidarity: *posture* and *locomotion*.

4.1 POSTURE

In the first place, certain centers in the brain *continuously* discharge impulses to the spinal cord to create a basal state of *readiness for action*, which is the physiological basis of normal posture. As with spinal reflex tonus, which contributes to normal posture, the basal motor state produced by the brain is the arithmetic sum of both excitatory and inhibitory impulses. The excitatory impulses originate from the greatest portion of the lower brain, in an area called the "bulboreticular facilitatory formation", which includes part of the medulla oblongata, all of the pons and mesencephalon, the thalamus, subthalamus and hypothalamus. These impulses travel to the spinal cord as well as to the forebrain or cerebral cortex. The mesencephalon in particular controls basal thalamic and cortical activity to produce either wakefulness or sleep, although, if any part of the facilitatory formation is not inhibited both muscle tonus and cerebral activity are greatly increased. In the normal animal, a region of the lower three-fourths of the medulla, called the "bulboreticular inhibitory formation", inhibits the facilitatory formation, although not directly. The basal ganglia and the cerebellum, which are considered parts of the lower brain, as well as the forebrain issue excitatory impulses to the inhibitory formation that enable the latter to exercise its inhibitory effect on the facilitatory formation. In addition, the basal ganglia, cerebellum and cerebrum directly inhibit the facilitatory formation, and thus the intrinsic excitatory nature of the latter is offset by these two inhibitory

processes. From this brief description it is apparent how, in an unmutilated animal, the excitatory and inhibitory impulses originating in various regions of the brain are pitted against one another to produce a basal state of readiness for action. These antagonistic processes provide the animal with sufficient nervous energy for it to maintain its balance and support its weight against gravity in a manner that defines its normal body posture.

By contrast, if either the basal ganglia, cerebellum, or cerebrum is removed the facilitatory formation exercises insufficiently inhibited influence over the spinal cord to produce a state Sherrington termed, "decerebrate rigidity". During such a state, all skeletal muscles, especially the extensors, are thrown into an extreme state of contraction characteristically causing the animal's limbs to stretch stiffly downward, and its head and tail to stretch rigidly upward. Sherrington also demonstrated the physiological counterpart to decerebrate rigidity, which he termed "spinal shock". A high spinal transection, which severs the cord from the influence of the facilitatory formation, causes a profound depression in the muscle tonus of all skeletal muscles aboral (tailward) to the transection. The depression renders the limbs incapable of reflex responses of any sort, for hours in a frog, to months or even years in a man. Thereafter, the neurones of the various reflex-arcs increase their intrinsic excitability, sometimes becoming hyperexcitable, and the capacity for reflex action returns. Sherrington attributed spinal shock to a "defect of transmission at the synapse" (IANS, 245) a defect that is due to, presumably, an absence of impulses from the brain that are normally available for the synapses to transmit. Sherrington

learned about the brain's contribution to posture by means of surgical experiments such as the above.

Although spinal reflex tonus contributes to normal body posture, there is a marked difference between the posture produced by the cord alone, and that produced when the brain is intact. And this pointed to the difference between motor coordination, per se, and higher order motor solidarity. The overall posture of a spinal animal is simply a concatenation of the postural effects produced by each spinal segment in the series. The stretch receptors in a given muscle inform a particular segment of a definite contractile state of the muscle, more in relation to its antagonist than in relation to the remaining limbs. In contrast, the labyrinthine receptors of the cranial segment inform the brain, especially the cerebellum, of the postural states produced by *every* segment of the body. Such information defines the animal's posture less in relation to individual segments, and more in relation to the external world. This enables the brain to respond in a manner that engenders motor solidarity, whereas a spinal segment has merely local jurisdiction. On this difference Sherrington writes:

> "A posture of the animal as a whole -- a total posture -- is as much a complex built up of postures of portions of the animal -- segmental postures -- as is the total movement of the animal -- its locomotion -- compounded of segmental movements. With the hinder part of its spinal cord alone intact the frog maintains a posture in its hind limbs. These limbs are kept flexed at hip, knee, and ankle. When displaced from that posture they return to it. But if the animal be rolled over onto its back it makes no attempt to right itself. The decerebrate frog with its labyrinths intact and their arcs still in connexion with the skeletal musculature maintains the well-known attitude before mentioned. If inverted it at once reverts to that. The labyrinth keeps the world right-side

> up for the organism by keeping the organism right-side up to its external world. The cranial receptors control the animal's *total* posture as do receptors of the hinder musculature the *segmental* posture of the hind limbs when but the hind end of the spinal cord remains." (IANS, 341)

Thus motor solidarity engendered by the brain is recognizable even in regard to posture, which can be considered the most primitive expression of motor solidarity.

4.2 LOCOMOTION

However, as the above passage suggests, there is not a sharp division between posture and locomotion, as Sherrington states in this passage:

> "Naturally, the distinction between reflexes of attitude and reflexes of movement is not in all cases sharp and abrupt. Between a short lasting attitude and a slowly progressing movement the difference is hardly more than one of degree. Moreover, each posture is introduced by a movement of assumption, and after each departure from posture, if it is resumed, it is reverted to by a movement of compensation. Hence the taxis of attitude must involve not only static reactions of tonic maintenance of contraction, but innervations which execute reinforcing movements and compensatory movements. In all this kind of function the proprioceptors (stretch receptors) of the body generally and the labyrinthine receptors in the head, appear to cooperate together and form functionally one receptive system." (IANS, 341)

Accordingly, the locomotive acts initiated by a unique class of receptors that feed into the brain, especially into the cerebral cortex, are characteristically expressive of motor solidarity. Sherrington called these receptors "distance-receptors" (the eyes, ears, and nose), because they react to objects at a distance. As we shall discuss

shortly, the reactions distance-receptors initiate have what Sherrington called a "precurrent relation" to subsequent reactions that gives the distance-receptors a singular biological importance. What needs immediate clarification is the solidarity of the motor acts these distance-receptors initiate; Sherrington writes:

> "As initiators of reflex movements the action of (the distance-receptors) is characterized by tendency to work or control the musculature of the animal as a *whole* -- as a single machine -- to impel locomotion or to cut it short by the assumption of some total posture, some attitude which involves steady posture not of one limb or one appendage alone, but of all, so as to maintain an attitude of the body as a whole. Take for instance, the flight of a moth toward a candle, the dash of a pike toward a minnow, and the tense steadiness of a frog about to seize an insect. These reactions are all of them excited by distance-receptors though in one case the musculature is impelled toward the stimulus (positive phototropism), in the other restrained (inhibited) from locomotion. Whether the reaction be movement toward or movement away from (positive or negative) or whether it be motion or its restraint (excito-motor or inhibito-motor) does not matter here. The point here is that in both reactions the skeletal musculature is treated practically as a *whole* and in a manner suitably anticipatory of a later event. (IANS, pp. 326-7)

Later in the passage, Sherrington points out that such holistic motor involvement is far less evident among responses evoked by exciting the exterio-receptors (non-distance-receptors) populating the receptive field of a spinal animal's hide; the latter receptors evoke merely local responses, as we saw previously. Thus, the difference between the holistic responses evoked by distance-receptors and the local responses evoked by non-distance-receptors is further evidence of the emergence of motor solidarity as a behavioural property that is qualitatively distinct from, although

physiologically dependent upon, motor coordination.

Behavioural reactions initiated by the eyes, ears and nose, the distance-receptors, in addition to being characteristically holistic, involve two stages. The immediate behaviour so evoked is best described as *anticipatory* to a subsequent *consummatory* behaviour. Since the subsequent behaviour is usually a reaction evoked by non-distance-receptors, the distance-receptors are said to have a "precurrent relation" to the non-distance-receptors. For example, consider the behaviour of a pike about to seize a minnow. Its visual, auditory and olfactory receptors initiate its wait, and then its dash toward the minnow. This behaviour precedes or anticipates a final or consummate reaction, i.e. of swallowing the minnow. However, the final reaction, whatever it may be, is determined by non-distance-receptors, most likely those populating the inner surface of the mouth. Should the object seized by the pike actually be a minnow, taste-receptors will most likely evoke the swallowing reflex. On the other hand, should the apparent minnow actually be a fisherman's lure, pain-receptors in the mouth will most likely inhibit the swallowing reflex, and the pike will probably react in an attempt to expel the lure. In either case, the reaction evoked by the distance-receptors remains the same, and is therefore independent of subsequent reactions evoked by the non-distance-receptors. By contrast, the reactions evoked by the non-distance-receptors are most unlikely to occur without the occurrence of the anticipatory behaviour evoked by the distance-receptors.

This last comment suggests the value of the distance-receptors to the animal and its species. As Sherrington writes:

> "The distance-receptors seem to have a peculiar importance for the construction and evolution of the nervous system. In the higher grades of the animal scale one part of the nervous system has, as Gaskell insists, evolved with singular constancy a dominant importance to the individual. That is the part which is called the brain. *The brain is always the part of the nervous system which is constructed upon and evolved upon the distance-receptors.* Their effector reactions and sensations are evidently of paramount importance in the functioning of the nervous system and of the individual. This seems explicable, at least partly, in the following manner.
>
> An organism is not a machine which merely transforms a quantum of energy given it in potential at the outset of its career. It has to replenish its potential energy by continued acquisition of suitable energy-containing material from the environment, and this material it has to incorporate in itself. Moreover, since death cuts short the career of the individual organism, the species has to be maintained, and for that in most higher organisms there is required accession of material (gametic) from another organism (of like species) to rejuvenesce a portion of the adult, which portion then cast off leads a new individual existence. To satisfy, therefore, the primary vital requirements of an animal species, actual material contact with certain objects is necessary; thus, for feeding, and in many cases for sexual reproduction.
>
> In these processes of feeding and conjugation the non-distance-receptors play an important and essential part. But ability on the part of the organism to react to an object when still distant from it allows an interval for preparatory reactive steps which can go far to influence the success of attempt either to obtain actual contact or to avoid actual contact with the object." (IANS, pp. 325-6)

To conclude my discussion of motor solidarity, I will sketch how the nervous system is organized to perform locomotive acts. Basically, ordinary locomotion requires that three main neural centers be integrated with each other: the forebrain, or cerebral

cortex, the lower brain, or cerebellum, and the spinal cord. The cerebellum is the intermediary organ integrating the forebrain, which issues motor orders, with the spinal cord, which translates the orders into actions. It functions by comparing actual contractile states of the skeletal muscles, as depicted by incoming sensory information from the proprioceptors and the labyrinth, with the outgoing motor directives issued to the motor system by the cerebral cortex. The cerebellum monitors and makes corrective changes in the motor activities as they occur, and is therefore of great importance to activities such as running, typing, piano-playing and even talking, where subtlety of adjustment in muscle contraction is imperative and rapid. Sherrington's research of the brain would be continued by his students who uncovered the details about the structures and functions necessary to motor solidarity, in respect to both posture and locomotion.

4.3 SUMMARY

Our investigation thus far has uncovered what has been considered to be three levels of reflex integration. Our findings can be summarized in Chart III. (See page 105)

Chart III shows that by means of the integrative action of the nervous system three levels of observable complexity emerge: motor coordination within a given reflex; motor coordination between reflexes, both allied and antagonistic; motor coordination augmented and regulated to effect motor solidarity. Moreover, in general, greater complexity entails greater dependence, especially upon 'simpler' conditions. For instance, motor solidarity depends upon the necessary conditions of motor coordination: the all-or-nothing

Chart III

INTEGRATIVE ACTION

level of behavioural sophistication	1st order coordination	2nd order coordination	motor solidarity
behavioural feature	response appropriate to stimulus	compounding of reflexes	posture locomotion
physiological properties	1. selective excitability 2. conduction to afferent 3. unidirectional conduction 4. threshold conduction	5. convergence to FCP 6. reciprocal innervation/ inhibition	7. lack of reflex neutrality 8. lack of reflex inevitability 9. increase in reflex precision
anatomical necessities	1. receptor organ 2. neurone/synapse 3. afferent neurone 4. efferent neurone 5. effector organ	a multiplicity of necessities 1-4, and an effector organ	a multiplicity of necessities 1-5, plus: 6. lower brain 7. midbrain 8. forebrain

behaviour of intercellular conduction; the irreversibility of intercellular conduction; fractionation and occlusion, the physiological products of converging afferent neurones to a final common pathway; and reciprocal inhibition. In addition, motor solidarity depends upon the distance receptors, and various parts of the brain. By contrast, 'lower level' motor coordination is producible without centers higher than the spinal cord. In short, Sherrington's integrative action theory shows how a multiplicity of behavioural complexities is created and welded together by the nervous system to produce the functional unity which characterizes the behaviour of a vertebrate animal.

5. PURPOSIVENESS IN THE REFLEX

Common to the behavioural diversity so produced, and perhaps indicative of how functional unity is attained is the striking feature that at all levels, *behaviour is purposive*. Purposiveness is an obvious aspect of mindful behaviour. If I go to the store, I go with a particular purpose; if I study, I study for a reason; if I invest, I have a goal in mind. Indeed, mindful behaviour that is not purposive is very often not intelligent. Equally, the object of a scratch reflex is neither indiscernible, nor easily confused with the purpose of, say, a thrust reflex. Sherrington asserted that such spinal behaviours are "not meaningless; they carry each of them an obvious meaning." (IANS, xi) Equally again, spinal behaviour is not intelligible unless it has a distinct purpose:

> The purpose of a reflex seems as legitimate and urgent an object for natural inquiry as the purpose of the colouring of an insect or a blossom. And the importance to physiology is, that the reflex reaction cannot be really

> intelligible to the physiologist until he knows its aim. (IANS, 237)

Here again is Sherrington's clear commitment to the belief that teleology is the proper domain of scientific physiology. Without teleology, knowledge in physiology is impossible, according to Sherrington. Yet, contrary to what modern philosophers such as Nagel hold, such a commitment to teleology does not further necessitate adopting the view that teleology is special causality. What is entailed in Sherrington's commitment to teleology is adoption of the view that *teleology denotes a special organization* of ordinary causal processes. Thus we find Sherrington in Lecture VII outlining the rules for assigning purposes to particular reflex acts, and these rules, as we shall see below, make reference to the nervous system's organization as characterized by what he termed first and second order coordination, respectively. In other words, in knowing the integrative characteristics of a given reflex, learned through repeated observation, one is able to inductively surmise the particular purpose of the reflex in question. Allow me to elaborate.

Sherrington developed a method to determine the purposes of different reflexes, the so-called "fractional pieces" of an animal's total behaviour. He was not concerned to inquire about the "intricate purposes ... traceable in these total reactions which constitute the creature's behaviour as a social unit in the natural economy ..." (IANS, 238) Rather, he wanted to demonstrate, by empirical analysis, that so and so reflex had such and so purpose. The method involved "repeated observation" of "one or other of the conditions attaching to the (reflex) reaction." (IANS, 238)

> The *mode* of adequate stimulus is one of these. The time-relations and spatial form of the response are others. ... and ... The accessory parts of (a reflex reaction) are often instructive concerning the whole. (IANS, 240)

The mode of adequate stimulus, the time-relations, and the spatial form are all conditions, we have seen, that characterize first order coordination. Thus, one must understand first order coordination, coordination that renders responses appropriate to stimuli, to determine the purpose of a reflex. Moreover, knowledge of the accessory parts, or of the particular coordination present between the main reflex under study and all others observable at the time, is also necessary to determining the purpose of the reflex under study. Thus second order coordination, coordination between reflexes, tells the observer something of the purpose of a reflex as well. Sherrington offers a helpful example:

> The broad pressure applied under the foot-pad which seems the adequate stimulus for the 'extensor-thrust' and the brief forcible straightening of the limb which constitutes the response suggest that the reflex has its purpose in the execution of an act in the series of movements of stepping in the animal's locomotion. ... Again, the connexion between the tickling irritative stimuli which seems 'adequate' for the 'scratch-reflex', and the scratching movement itself which results, suggests that the purpose of that reflex is a grooming of the skin to protect that organ against parasites which infest it and would confuse its function as a receptive surface reacting to more significant environmental stimuli. (IANS 239)

The modes of adequate stimuli for the extensor thrust and scratch reflexes are, respectively, pressing and tickling, their respective time-relations are, aphasic and phasic; and their respective spatial forms may be described as, straightening of the limb, and drawing-in then reaching with the limb. On the fourth type of datum, concerning the accessory parts, in the case of the scratch reflex, Sherrington observed:

> The accessory parts of the reflex, namely those in the three limbs which are not scratching, are also contributory to the same effect as in the scratching movement in the right hind limb itself. They steady the dog and secure the stability of its body during the performance of the scalptor act. (IANS, 240)

Lastly, it thus appears that for Sherrington, assigning a particular purpose to a reflex is not unlike generating an ordinary empirical hypothesis. From a number of controlled observations, experiments, a set of data is collected. Routines and regularities are determined from an analysis of the data, and relationships are surmised; from the collected data, one reasons inductively as to the purpose of a given reflex. Presumably, the certainty about whether such and so is the inductively derived purpose of so and so reflex would depend upon an absence of particular instances of identical behaviour having contrary purposes. Scratching, for example, would not be the purpose of the scratch reflex if it were learned that the characteristic motion always failed to remove the irritant, or that the 'irritant' proved not to be irritating, or if some effect other than removing the apparent irritant was discovered.

In short, by observing reflex behaviour, through the light of knowledge of the integrative action of that behaviour, it was possible for Sherrington to inductively derive the purposes of various reflexes.

Assigning purposes to reflexes, teleology, is justified, Sherrington held, owing to Darwin's theory of evolution. Certainly, in Sherrington's view, Darwinism had dissolved the age-old problem about whether mental events accompany spinal reflex reactions. Darwinism showed how a spinal event could be both purposive yet

non-mental:

> (Darwinism) suggests how purposive neural mechanisms may arise. It furnishes the key to the gensis and development of adapted reactions and, among these latter, reflexes. That a reflex should exhibit purpose is no longer considered evidence that a psychical process attaches to it; let alone that it represents any dictate of 'choice' or 'will'. In light of the Darwinian theory every reflex *must* be purposive. We here trench upon a kind of teleology. (IANS, 236)

Thus a purposive act is an adapted reaction. But Sherrington does not enter into a discussion of how Darwinism justifies the use of terms such as "goal", "purpose" or "aim" in the context of discussions of non-intelligent spinal reflex reactions. His attention turned instead to the experimental difficulties the phenomenon of spinal shock posed for his effort to discover the purposes of various reflexes. My discussion will now depart from his, and focus first upon the question of the Darwinian justification that Sherrington probably had in mind.

A main mechanism of evolution is the process of natural selection, which changes the phenotypic properties of a given population over time. Darwin observed individual variation among members of the same species, and surmised that some variants would have an advantage over other members in the struggle for life. The favourable variants would transmit their advantageous properties to their progeny, and since populations produce many more progeny than what the environment will permit, the proportion of favourable variants which survive to reproduce will be larger than the unfavourable. Thus there will be continuous evolutionary change, producing new variants, species, genera, and so on.

Sherrington believed that a purposive reflex act, such as scratching to remove a flea, is a process that developed over generations because of a more or less constant element in the species' environment e.g. fleas (irritants). Individuals having the capacity for scratching and grooming their hides were perhaps less susceptible to disease, generally "better able to do dominate their environment", and thus more likely to have progeny. These, in turn, would have to compete against a population whose greater number is so adapted in respect to a scratch reflex. If the environment continued to selectively favour such a purposive reflex act, the fittest among the offspring will be, *ceteris paribus*, those whose scratch reflex is most adaptive, i.e. purposive.

In equating 'purpose' with 'adapted reaction' has Sherrington settled one philosophical problem, only to overlook another? Darwinism does enable 'purposiveness' to be explicated without mentalistic phraseology, but it does not reduce teleological terms to causal ones. For evolutionary theory does not characterize an individual purposive act *qua* individual act, let alone explain the causality of the individual act. At most, the theory hints at the causality of purposive behaviour by rendering the striking sophistication an end result of an ordinary, forward-going developmental process. To explain *individual* purposive acts we must rely upon the theory of integrative action, not evolution. Yet in so doing, we of course encounter unreduced teleological concepts: recall, "the reflex reaction cannot be really intelligible to the physiologist until he knows its aim." (237) Can such inductively derived teleological concepts be reduced to causal ones, or are they a necessary and

fundamental aspect of physiology, as Sherrington claims? Nagel's model of self regulation can now be tested against Sherrington's theory of integrative action to ensure its generality, and against Sherrington's contention that teleology is to be used, not avoided.

CHAPTER III

SHERRINGTON'S ANSWER TO NAGEL

I can now give critical scrutiny to Nagel's model of goal-directed behaviour in view of Sherrington's theory of the integrative action of the nervous system, a goal-directed system *par excellence*. Recall that for Nagel, a purposive or goal-directed system can be reduced to three primary properties of the behaviour of such a system; the behaviour was claimed to be: plastic, persistent, and orthogonal. These properties, in turn, are further reducible to, or solely attributable to, the interplay between the parts of the system, insofar as any primary variation in the state of a part is offset by some adaptive variation in the remaining parts, which returns the system to its homeostatic state, and gives its behaviour the above-mentioned primary properties. Clearly, therefore, showing that these properties are not truly predicated of purposive behaviour is a serious blow against the legitimacy of Nagel's reductionist efforts.

In the first place, Nagel's criterion of plasticity must be rejected: a process can be goal-directed, yet have but one pathway to the goal. Witness the mouse in a maze that has only one route to the cheese. The mouse's behaviour is clearly goal-directed even though it is restricted to a single course to bring about the result. Thus, goal-directed behaviour is not necessarily plastic. Moreover, the nerve impulses of a spinal reflex, say the scratch reflex, converge to what Sherrington called the final common path which leads to an effector organ, in this case the hind limb. An absence of alternative nerve paths is essential for coordination in the simple reflex, and equally vital in the compounding of reflexes. Were it not for a

final common path, antagonistic reflexes could send conflicting 'orders' to the same effector organ, and the animal would have no automatic means for controlling stimuli and coordinating responses. The spinal reflex represents a large class of obviously purposive processes that not only do not, but in fact cannot, evince the property of plasticity if their goal-directed organization is to be preserved.

Secondly, we must qualify his criterion of persistence. It would be wrong, however, to argue against Nagel that, persistence is not a property of goal-directed systems because these systems are not always able to persist in their behaviour to the realization of the goal. This line of argument could cite the example of the individual who exercises to avoid heart disease, but who dies of coronary thrombosis while jogging. Again, Sherrington's study of antagonistic reflexes showed that certain reflexes have pre-emptive ability over others, and thus the impulses from an ongoing reflex may be blocked by a second reflex put into action. This criticism of Nagel will not stand however, because it is necessary to assume that the system is in fact persistent to argue that the breaking of persistence, either by the death of the system or by the intervention of a mechanism aiming at a higher level goal, proves that persistence is not a true property of goal-directed systems. In short, the argument supported by the examples presupposes what it attempts to disprove , and thus it can't succeed. Although failure to realize a goal does not affect the claim that purposive behaviour is persistent, failure to realize adaptive variations requires that the claim be qualified. A system may be purposive, yet never show the property of persistence. That is, it may never be called upon to make adaptive

changes, for the internal and external conditions may be such that no disturbance ever occurs. Thus, should we use the criterion of persistence consistently when attempting to identify goal-directed systems, as Nagel would have us do, we would overlook some systems which are not actually persistent, but which are in fact goal-directed. At best, the criterion of persistence should be qualified to become the criterion of potential persistence: goal-directed systems must have the capacity to be persistent, although they may never have to be persistent.

Third, orthogonality as characterized by Nagel is misleading and causes us to overlook an important feature of goal-directed systems. The variables (A, B, and C) are logically independent, but *physiologically interdependent*. Logical dependence would dictate, as previously indicated, that the state of one variable could be logically deduced from knowledge of the state of the more primitive variable, and it would thus be redundant to mention the former in defining the primary conditions of the system at hand. However, it should be pointed out that a deduction of the sort Nagel wants to guard against is possible only if our question about the state of any particular variable can be raised in the context of a knowledge of the state of the entire set of remaining conditions, both internal and external. It is impossible, for example, to deduce that the left hind limb of an animal is scratching, and not withdrawing, from information about the state of what Sherrington called "the accessory parts" alone, namely, the remaining three limbs. It would seem that a deduction of this short is possible if we know at least something of the reflex in question, e.g., that the tactile receptors of the

hide are being stimulated, and not the pain receptors of the left foot. Conversely, from knowing that the scratch reflex is operating, it does not follow that the accessory parts are in one particular state and not another, e.g., positioned for standing, as opposed to being relaxed with the animal sitting for a scratch. In short, deducing the state of one variable from that of another is not possible with incomplete information, and thus Nagel did not have to worry that the logic of his model of a goal-directed system would collapse into a description of one and only one of its parts. What he intended, of course, was to protect the reasonable distinction between purposive systems and ordinary systems of equilibrium, without relying upon the concept of a goal to do the job. But if my criticism is sound, the difference between the two for Nagel is a function of the completeness of our information about each system: we treat closed systems in defining equilibrium, but epistemologically open, or partially defined, systems when defining goal-directed behaviour. Clearly, such arbitrariness renders the distinction between equilibrium and goal-direction one of difference in degree, not kind. To say that goal-directed systems are orthogonal is not to distinguish them from equilibrium systems, *per se*, and this is unacceptable.

Moreover, Nagel's fruitless argument for logical independence is misleading in that it causes us to overlook the physiological interdependence between parts in a goal-directed system. As Sherrington demonstrated, the state of any one component of such a system in some way shapes and determines the state of the remaining internal conditions. For example, the accessory parts to a scratch reflex, the three remaining limbs, can facilitate the scratching motion by steadying

the animal. Conversely, the scratch reflex of one limb will itself influence the state of the remaining limbs for the duration of its dominance. State variable interdependence is a crucial feature in the compounding of reflexes, as Sherrington has shown, and thus a necessary condition of coordination and purposive behaviour. Furthermore, it is precisely because of the causal interdependence among variables of a directively organized system that we exclude the likes of rivers from being purposive processes: although the direction a river takes is causally dependent on the lie of the land, at some time t_1, the lie of the land is not causally dependent upon the river flow at the same moment. On the other hand, the body core temperature will influence the diameter of the peripheral vessels, and *vice versa*. Even the position of a flea will affect the state of a hind limb which, in turn, may significantly alter the state of the flea. So the variables must be physiologically interdependent, though they may be logically independent. A close study of the integrative action of the nervous system -- a goal-directed system if there ever was one -- leads us to reject all three of Nagel's criteria for goal directedness, and so upsets his reductive proposal.

The heart of the trouble with Nagel's theory of directive organization stems from the very attempt to avoid teleological terminology, yet characterize the essential concept of a teleological system. The effort fails. First, the concept of orthogonality can safeguard the distinction between goal-directed and equilibrium systems provided its meaning entails the concept of a goal. Whereas the formal meaning of "orthogonality" which Nagel defines is fruitless and misleading, an informal definition, in the sense of 'self-

straightening' or 'self-correcting', can secure the distinction between these two kinds of systems. Goal-directed systems are self-straightening, orthogonal, with respect to some goal. Equilibrium systems, on the other hand, are not orthogonal because it is empirically meaningless to refer to the "goal" of such a system. Second, recall that for Nagel primary variations remove the system from its G-state, and adaptive variations, or compensatory changes, return the system. The analysis is covertly teleological! An adaptive variation is an adjustment to a change; but not just *any* adjustment can be considered truly adaptive. Such adjustments are adaptive with respect to some *future*, yet-to-be-achieved *goal*, i.e., with respect to a condition that is not presently predicated of the system, but will be so predicated at a future time, *ceteris paribus*. Directive-organization cannot be described without tacit reliance on the concept of a goal; such organization is inherently teleological. Third, purposive behaviour is *recognizable* only if it is possible to ascribe a goal to the behaviour in question. At most, Nagel's three distinct characteristics are clues that *may* alert the observer to the presence of a goal-directed system. However, the absence of any one or all three of these conditions should not lead the observer to conclude that what is before him is not a goal-directed system. Even when these properties *are* present, what clinches the case is whether the properties exist with respect to a goal. If a goal is not discernible, the properties do not describe a goal directed system. The reason the knee jerk, *qua* knee jerk, is not a goal-directed process, is not because it lacks persistence, as Nagel asserts (270, "Teleology"). A knee jerk has no purpose in the animal economy, and therefore it

is not a teleological event. It is, however, *indicative* of a goal-directed process, namely, the tonic reflexes of the limb in question. In the clinic, where a knee jerk is evoked for the purpose of causing muscle tonus and gaining insight into neurological disorder, the knee jerk becomes a teleological event.

For Sherrington, purposiveness is as much a quality of reflex action as is the presence of a receptor organ. Knowledge of reflex action would be as incomplete without teleology as it would be without the notion of a receptor organ. Sherrington's cause-physical account of reflex action begins with a receptor organ, leads us from afferent neurone to grey matter, to efferent neurone, to effector organ and terminates with the causal outcome, a coordinated, purposive act. Both receptor organ and purposive act are observable and necessary to an explanation of the integrative action of a spinal reflex. Both concepts are equally susceptible to becoming emptied of empirical meaning, if taken from the context of the integrative action theory. 'Receptor organ' is meaningless without the concepts of 'neurone/synapse', 'afferent' and 'efferent arcs', 'grey matter', 'motor coordination' and so on; i.e., without the theory of the integrative action of reflex behaviour. Equally, purposive reflex behaviour is either a misleading or an empty phenomenon without the context of the theory of the integrative action. Yet purposive behaviour is not merely an effect; it is an effect which can, in turn, *cause changes* in the system itself. It is not an epiphenomenal product; it can be a causally efficacious *control mechanism*. Purposive behaviour can be studied, as Sherrington has shown, and its effects on the teleological system can be measured as modern physiologists are doing. In regard

to causal efficacy as well, then, purposive behaviour enjoys the same status as any anatomical feature necessary to reflex action.

By reaffirming the place of teleology in the study of organic systems, Sherrington anticipated the primary interest of present day physiology.

> Until recently, the principles of control systems as applied to the human body have been applied only qualitatively rather than quantitatively. For instance, in the case of baroreceptor regulation of arterial pressure, physiologists have simply stated that an increase in pressure causes a reflex decrease in pressure back to normal. However, this means very little unless we know *how much* effect occurs. But within the past 15 years, application of much more quantitative principles of control systems has begun to make physiology a much more exact science than it has been in the past. (Guyton, pp. 7-8)

With the above in mind, consider the following example from Guyton of a purposive process, the homeostatic control of temperature. (See page 122).

The physiological interest in mechanisms of control focuses squarely upon purposive behaviour, *qua* purposive behaviour. The causal efficacy of the system upon itself, of these purposive mechanisms, is the primary interest. They are not regarded as epiphenomenal and understandable only in terms of some underlying microprocess, as Nagel would have us believe. Purposiveness is an emergent property of the material, an organizational property of matter, just like mentality. Purposiveness, as a mechanism of control and management, in and of itself, is as much the rightful domain of scientific physiology as mind is of psychology. Nagel has actually misconstrued the problem: teleology need not be considered a special causality, as it once was, but a special organization of

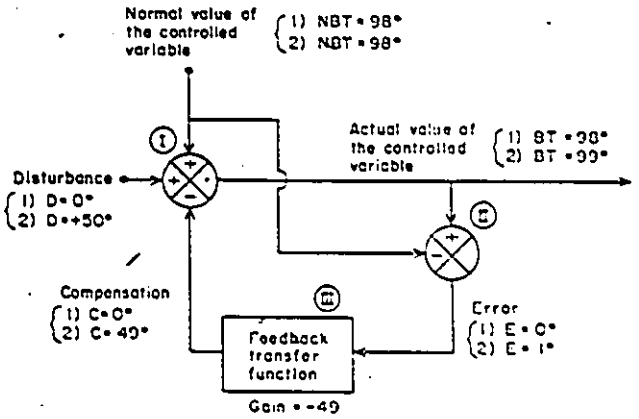


Figure 1–5. Static analysis of a control system. This figure, presenting an analysis of the control of body temperature, shows values for different variables in the system under two different conditions: (1) normal, and (2) when there is an increase in air temperature of 50° F, thereby creating a disturbance that attempts to raise the body temperature. (NBT = Normal body temperature; BT = Actual body temperature; E = Error; C = Compensation; D = Disturbance.)

tual value of the controlled variable and the normal value. The output of this block is called the *error*.

Block III is the feedback portion of the control system that responds to the error and determines the degree of compensation that will occur to overcome the disturbance.

Application of the General Analysis to the Control of Body Temperature. Now let us apply this general analysis of a control system to an actual example, the control of body temperature. At each point in the system two different sets of values are shown. The first value in each instance represents the normal when there is no disturbance and no compensation by the control system. Thus, the normal values are as follows:

Normal value for the controlled variable 98°(F.)
Disturbance ..... 0°
Actual value of the controlled variable.. 98°
Error ..... 0°
Compensation ..... 0°

Next, let us add a disturbance that attempts to change the body temperature. In this instance, we will assume that the air temperature increases by 50°. If the body temperature control system were not functioning, this would increase the body temperature by 50°. However, as the body temperature begins to rise, the actual temperature becomes different from the normal value, an error develops, and the control system provides sufficient compensation to overcome most of the disturbing effects of the abnormal air temperature. This compensation results from evaporation of the sweat from the surface of the skin, which in turn causes increased heat loss from the body. Therefore, the second set of values shown in Figure 1–5 is the following:

Normal value of the controlled variable 98°(F.)
Disturbance ..... +50°
Actual value of the controlled variable.. 99°
Error ..... 1°
Compensation ..... −49°

Thus, one sees that after the control system has become effective, the actual value of the controlled variable becomes equal to the sum of three different values: (1) the normal value for the controlled variable (98°), (2) the disturbance (+50°), and (3) the compensation (−49°), giving an actual value of the controlled variable of 99°. That is, the abnormal disturbance of +50° causes a rise in body temperature of only 1°.

One can then calculate the gain of this system by dividing the compensation by the error (−49° divided by 1°), yielding a feedback gain of −49.

Figure 1–6 illustrates the effect of this 50° increase in air temperature under two conditions: (1) when the temperature control system is not functional, and (2) when it is functional. When the system is functioning, the compensation increases at the same time that the body temperature rises. Therefore, by the time the body temperature has risen from 98° to 99° the control system will have already initiated −49° of compensation. The body temperature, therefore, will rise only to 99° rather than to 148°, an increase that would occur if the control system did not exist.

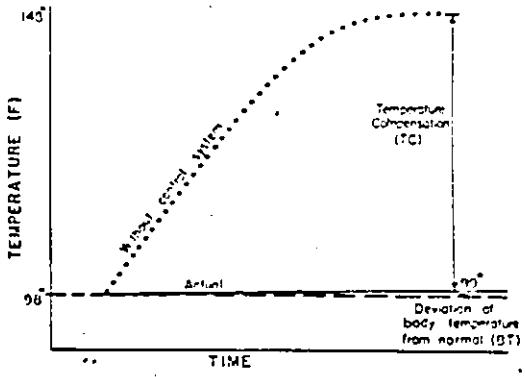


Figure 1–6. Effects on body temperature of suddenly increasing the air temperature 50° F, showing the hypothetical effect without a control system and the actual effect with a normal control system.

ordinary causal processes. The causal efficacy of a purposive system relates directly to its special organization, and such organization has become an area of inquiry and quantifiable knowledge in physiology.

CHAPTER IV

SHERRINGTON'S ANSWER TO HEMPEL

THE RESOLUTION OF HEMPEL'S TRILEMMA

Recall Hempel's conclusion that functional explanations, which argue from the observation of some recurrent activity to the presence of certain items functionally characterized, from effect to cause, are either invalid, nonempirical, or else trivial. Functional explanation is invalid, because although an item may have a function, the occurrence of the function does not imply the existence of the item; it becomes nonempirical if we attempt to save the deducibility criterion, in the only way possible, by defining the item as functionally indispensable; to save both the deducibility and empirical meaning criteria, we must render the explanation a deduction of a class of functional items I, but this will not provide for the explanation of particular functional items. Deduction of classes, not particulars, though valid and meaningful is rather "trivial", according to Hempel.

The trilemma can be resolved if the deduction of classes can be shown to have greater bearing than what Hempel admits. Helpful here is G.A. Cohen's remark that "it is no more a condition of sound functional explanation than of sound explanation generally that it explain the phenomenon in some ordained measure of specificity."

If we reconsider the role of functional explanation in physiology, in light of Sherrington's research and findings, the trilemma can be resolved. Let me begin with a brief general statement about the relation between the sciences of anatomy and physiology, as Sherrington appears to have developed that relation.

Anatomy and physiology are epistemologically symbiotic. Anatomical knowledge is logically necessary to physiological explanation. Reciprocally, physiological inquiry often reveals gaps in anatomical understanding and shows areas where further anatomical exploration is needed. For example, Sherrington knew that reflex action is a coordinated, purposive activity - it was the physiological problem he set out to solve - before the discovery of the anatomical synapse. His detailed study of the physiological process, it will be remembered, revealed eleven primary differences between nerve-trunk and reflex-arc conduction (see Chapter II, page 74). Obviously, reasoned Sherrington, there must be an anatomical basis for these eleven physiological differences; but owing to inadequate histological methods, Sherrington could not observe precisely what this basis was. Nonetheless, the overwhelming weight of physiological evidence forced him to propose that reflex-arc conduction is interneuronal, in addition to being intraneuronal; i.e., he conceived the idea of a synaptic gap between anatomically separate neurones in the reflex-arc. Once the presence of the synapse was proposed, he could begin to explain first order coordination as a physiological process. Much later, anatomists would observe the synaptic gap between neurones in a reflex-arc. A second example of reasoning from known physiological characteristics to unknown anatomical structures involves two steps. The principle of convergence was based upon the known anatomical situation wherein afferent-arcs outnumbered efferents 5 to 1. The principle explained the known physiological phenomena of fractionation and occlusion, which Sherrington demonstrated by comparing motor with reflex responses in the *Tibialis Anticus* muscle (see Chapter II,

page 82). In turn, however, fractionation and occlusion implied for Sherrington the unobservable anatomical situation he termed a "moto-neurone pool". Again, much later, the development of microtechniques enabled anatomists to confirm the anatomical situation Sherrington had proposed, based upon his knowledge of physiology.

What these examples show is that physiological inquiry *starts* with the assumption that certain, perhaps unknown, anatomical structures have functional significance. That is, certain structures are present, and these, in some way, causally contribute to the observed physiological effect. Functional/physiological inquiry is satisfied once the causally significant anatomical structures are known, for then the latter can enter into a causal/physiological explanation of the initially observed, and perhaps not understood, functional activity.

It does not follow from this that functional explanation can and should be replaced with causal explanation. I believe the value of functional explanation is in the fact it shows the functional significance, or causal importance, of the anatomical structures that comprise a physiological system. The following deductive schema for functional explanation is a resolution of the trilemma, for it shows how the deduction of a class of functional items is the duty of functional explanation in anatomical and physiological research.

(10a) At t, s functions adequately in a setting of kind c (characterized by specific internal and external conditions).

(b) s functions adequately in a setting of kind c only if a certain necessary condition, n, is satisfied. (e.g. blood-circulation)

(c) n is characterized by such and so features, and such and so are partly satisfied by an I-trait. (e.g. a pump of some form)

(d) Hence, at t, some I-trait is present in s. (follows from the fact of s's functioning adequately)

(e) Anatomical study shows i to be present in s, and i is an I-trait.

(f) Hence it is i that satisfies n that is necessary to satisfy s's adequate functioning.

From the above it can be said, a complete functional explanation (10a-f) points out the physiological significance of a *known* anatomical structure; whereas, an incomplete functional explanation (10a-d) shows where further anatomical research is needed. The aim is to deduce physiological or causal significance, not anatomical presence, when anatomical knowledge is complete. The account sits well with what I have learned of Sherrington's research methods, and it meets the three formal criteria of Hempel's D-N explanation.

CHAPTER V

CONCLUSION ON THE DEBATE ABOUT TELEOLOGY

Organic activities exemplify the same laws which govern non-living systems. Teleology is not a special, occult, mode of causation, as was once thought, but a special organization of ordinary causal processes. The organization of a teleological system is considered special because it provides for the causal efficacy of the system upon itself. That is, unlike ordinary causal processes, the effects produced in a teleological system have the unique property of being themselves causally efficacious upon their causal source. Physiologists call the special organization which produces this capacity of a teleological system a "feedback loop". The feedback capability of a teleological system distinguishes it from a system of simple equilibrium. The thermal equilibrium which will obtain between the temperature of an open container of water placed inside a closed container of air is the result of an ordinary causal process. If the water temperature is greater than the ambient temperature, the cooling process the water undergoes has two effects: it will raise the ambient temperature while lowering its own temperature. The raised ambient temperature does not cause the water temperature to lower, nor can it cause the water temperature to increase, *ceteris paribus*. In contrast, the feedback mechanism of the sugar-insulin system, for example, distinguishes it from a situation of simple equilibrium. Sugar regulation has at least three causal steps: (1) a raised insulin level causes (2) a lowered sugar level, and a lowered sugar level causes (3) a lowered insulin level. Equally, in reverse: (1) a raised sugar level causes (2) a raised insulin level, and this in turn causes (3) the sugar level to drop back to normal.

The reciprocal relation between the components of a feedback mechanism, or loop, can be such that either the initiating cause is reinforced by its effects or, as with most cases of biofeedback, and as with the sugar insulin example, the initiating cause produces an effect which diminishes the potency of the initiating cause. The first type of teleological system incorporates a positive-feedback mechanism, whereas the second type is the more common negative feedback mechanism. Positive feedback is believed to be responsible for muscle contraction during childbirth: as the child exits through the vaginal canal, it stretches the vaginal muscles which causes them to contract; as contraction of these muscles occurs, the child is pushed through the canal causing further stretch-receptors to become stimulated causing greater contraction and so on until the child makes good his exit. Physiological drug addiction is another example of the positive feedback phenomenon: the more a given drug is consumed, the more becomes necessary before satiation is achieved. Negative feedback, on the other hand, is more typical of the biological process and is responsible for controlling the levels of various chemicals in the body, and the activities of various biological processes. Negative feedback is what is usually referred to as a control mechanism, since it is the mechanism by which homeostasis is achieved.

The special organization of biological processes that gives rise to positive and negative feedback loops requires special teleological categories and concepts such as 'goal', 'purpose', 'function', 'aim', etc. for its explication. The maintenance of a restricted range of values for certain internal conditions of the teleological system is what is referred to as homeostasis. The self-correcting,

or orthogonal, mechanism necessary to homeostasis is really only intelligible if teleological concepts and categories are used in any analysis. That is, although a homeostatic state is the result of past and/or present conditions within the system, the full meaning of what is occurring within the system at a given time is understandable only if reference is made to what the system is attempting to accomplish, and most likely will accomplish at some future time. Reference has to be made to a goal which lies in the future. This is what Sherrington discovered about the teleological process he encountered in the spinal reflex he studied: that understanding what the system is doing entails knowing what the system is attempting to do. Moreover, this is also what we discovered in Nagel's effort to describe, non-teleologically, the causal interplay between parts in a feedback loop: that his concept of an 'adaptive variation' tacitly refers to some unspecified notion of what the system is attempting to do. Undeniably, biofeedback mechanisms operate through ordinary causality. The scratch reflex, for example, which offsets increased neuronal activity in a local area, is produced by causo-physical processes occuring in the reflex circuitry involved and, in turn, the reflex scatching removes the irritant in a manner which is straightforwardly mechanical. There is nothing occult about any such example of purposive behavior.

The point is, however, that an adequate understanding of biological organization, at least in regard to control mechanisms, is inherently valuational. If we really hope to know what goes on in an organism, we must be prepared to assess the significance of what we see, the teleological significance certain parts and processes have for the organism. An organism is organized in a way which enables

it to satisfy certain requirements, such as growth and reproduction. Organisms and their subtle processes can pursue goals, just as we humans can. Whereas with psychology, the justification for using teleological language in accounts of human behavior ordinarily appeals to the organization of the subject's thoughts, beliefs and desires, for example, in physiology, the use of teleology is justified by appealing to the physiological organization of the system in question: the system has the organizational capacity to act purposefully. In short, what justifies the use of teleology in every case, physiological or not, is the degree of organization of the system at hand. In physiology, its use does not require any occult assumption about intelligence on the part of the system, as was supposed by early writers. Conversely, failure to use teleological language in physiology when one should, and sole reliance upon causal phraseology to explain purposive behavior, results in either tacitly permitting teleology into the analysis, as Nagel has done, or else failing to explain the significance of the physiological events we encounter The latter is tantamount to failing to explain the events at all.

From the above, it is apparent that teleological and causal analyses do not compete with each other but address different aspects of, in this case, biological organization. Through causal analysis, evolutionary theory leads us to understand how purposive mechanisms arose, as Sherrington intimated. Causal analysis is also necessary to understand the feedback mechanism which makes purposive behavior possible. Thus causal analysis is necessary to understand both how biological organization arose, and how it works. Teleological analysis, on the other hand, is necessary to understand the significance of the causal processes which compose the organization. The significance of

these is understood in relation to the goals of the system, such as growth and reproduction, as mentioned, and homeostasis, more immediately.

It should also be clear that teleological analysis presupposes a causal one, for our ability to discern the goals of a system depends on, in large part, what we know of how the system works. Sherrington demonstrated this point in developing rules for the assignment of purposes to various reflexes, which rules are based on a knowledge of the causality of the spinal reflex. Moreover, it also appears that once we know how a system works we know its goals, that in knowing the cause we know the purpose. Is teleology therefore, after all, banishable in principle? It is interesting in this regard to recall Sherrington's reflection on the history of his subject that, once the anthropomorphic image of an informing spirit which explained purposiveness was unseated, physiology became for a time "extremely reticent about purpose, remaining simply objectively descriptive". (IANS, p.237) The implication is of course that explanation is possible only if teleological terminology is admitted into accounts of purposive behavior. What I imagine Sherrington had in mind is that purely causal accounts are "simply objectively descriptive", since these only allow one to trace a chain of events leading, for example, from a receptive surface to the contraction of muscle in an effector organ. The overall structure of what is happening is lost by concentrating on the subtler processes which compose the event; one loses sight of the forest for the trees, as it were, unless teleology is relied upon in explanation. Philosophers such as Hempel and Nagel would reply to this, however, that any "overall structure" is causal, not teleological, and that the event in question is explained by subsuming each subtle process under an

appropriate causal law which would enable a deduction (explanation) of the outcome from the initial and boundary conditions of the system. Of course, as would be argued further, teleological terminology, such as "purpose", "function" and "goal" can be translated into language which is purely causal. A strictly causal account may prove cumbersome, and may yet be awaited, but with time, physiologists will develop and perfect a causal language more suited to the complex causality of biologically organized systems. In the mean time, teleology is a convenient shorthand account of a longer causal explanation; but convenience should not be confused with indispensability.

Is this conclusion of such philosophers really warranted? Could Nagel, for example, substitute a purely causal term for "adaptive variation" and still convey the meaning of the causal interplay between components in a biofeedback mechanism? Or will he always, necessarily, run into the trouble of tacitly relying upon some unspecified teleological notion? I reply 'no' to the first two questions, and answer the third in the affirmative.

I believe there is truth in what Sherrington supposed about the value of teleology, for although the organization of any individual process in a spinal reflex is clearly causal, the combined organization cannot be expressed non-teleologically. Moreover, I don't believe Nagel can avoid teleology when characterizing the purposive behavior of a feedback mechanism. The concept of an 'adaptive variation' refers 'back' to a primary variation, but also refers 'forward' to the likely outcome of the former's interaction with the primary variation, in so far as the notion is that the variation is truly <u>adaptive</u> or truly <u>compensatory</u>. In other words, one does not simply determine the number and kinds of steps in a causal feedback circuit,

however many causal laws there are under which each may be subsumed, but we make a further assessment about the end to which these events are heading: we must assess what the system is attempting to do, to understand what the system is doing. I am really not able to express the matter more clearly, other than to conclude that, in regard to the causal analysis of teleological systems, it too presupposes an, perhaps unspecified, account, namely a teleological one. Teleology and causality are thus different points of view of the same phenomenon, yet each is dependent upon the other and both are indispensable. With causal analysis, reasoning proceeds from cause to effect, with teleological analysis, the logic gets reversed; proceeding from effect to cause. The problem then becomes one of demonstrating the validity of such reasoning from effect to cause, which brings my conclusion to the subject of functional explanation.

Teleological, or functional, explanation ascribes a function or purpose to an item within a teleological system and argues from the occurrence of the function to the presence of some kind of item functionally characterized. That is, functional explanation, properly characterized, does not deduce the existence of particular items, but only that certain kinds of items are necessary to the recurrent or functional behavior observed. Thus from blood-circulation is deduced, for example, the presence of some form of pump; from reflex action is deduced, for example, the presence of some medium of communication. If we know that a system is functioning adequately at a given time; if we know that it does so only if certain processes obtain; and if we know that these processes are possible only if certain kinds of items are present, then we can legitimately deduce that certain kinds of items will be present. If this is all we know, then we must set about to

discover what kind of item actually exists which functions as a medium of communication, for example. Once the discovery is made, the function of the actual item becomes known, and this is what is meant by determining its significance: discovering how the part causally contributes to the observed recurrent or teleological behavior.

It might be challenged that with the discovery of which items are functionally significant goes the value of functional explanation: the latter is merely a dispensable hand-maiden to causal explanation. That is, although functional explanation indicates which items do what, and this appears to differ from causal explanation, which shows how which items do what they do, knowing-how entails knowing-what, and therefore teleological explanation collapses into causal explanation. Conceivably, as the argument would continue, functional explanation could be dispensed with entirely if we enjoyed a situation of perfect (i.e. complete) knowledge of biology. A rebuttal might be that, if anatomical discoveries remain, and if physiology can lead us to them, functional explanation is invaluable. On the other hand, it appears to be true that once we discover the structures responsible for particular causal networks in physiology, we secure a type of perfect knowledge, at least relative to a particular problem. Functional explanation does not add anything to our perfect knowledge of a given causal process, and in principle, therefore, it can be dispensed with. At most, as the argument would conclude, functional explanation is just a tool for obtaining knowledge, and for arriving at a point in our understanding where causal explanation becomes possible. It does not represent some additional understanding above and beyond the understanding represented in causal explanation.

The argument has great force, but misses the mark. The problem with functional explanation concerns whether reasoning from effect to cause is legitimate. I believe I have shown that it is legitimate reasoning. Conceivably, no matter how complete our knowledge of biology becomes, the biologist will want to use functional explanation to explicate the necessary conditions of a given biological system, and point to the significant roles of various structures which produce the effect he observes. Surely, it should be up to the biologist to decide how a biological problem is to be approached, and we, as philosophers, should be satisfied having determined the logical soundness of any approach he takes. Functional explanation is absolutely indispensable as long as discoveries remain, and as long as there is a physiological point of view.

To conclude, teleological organization is special by virtue of the causality of the feedback mechanism which produces purposive behavior. Special concepts and categories, teleological ones, are needed to understand the significance of any causal interaction occurring within such an organized system. Special explanation, functional explanation, is possible since such systems have the capacity to achieve nameable and observable goals. These goals become a viewpoint for analysis and explanation. Teleology cannot be reduced to causality, but the two are mutually dependent; together, they enable biological organization to be explained sufficiently. Teleology is compatible with materialism, the metaphysical assumption of science: as a phenomenon, teleology is a special type of causo-physical organization; as explanation, it presupposes that the organization, not the causation, is special, and proceeds from this assumption to explain the teleological significance, not the causation, of various items and processes, causes,

functionally characterized. Physiology cannot part with teleology.

APPENDIX

WAS SHERRINGTON A MIND-BODY DUALIST?

In Chapter II , we explored three levels of integration, first order motor coordination, second order motor coordination and motor solidarity, all of which Sherrington explained in terms of his physiological theory of the integrative action of the nervous system. Sherrington pursued the theory of integration further, applying it to perception and the self, and to understanding the relation between mind and body. Psychical solidarity and psycho-physical solidarity are, respectively, the fourth and fifth levels of integration which Sherrington discussed and explained by means of the integration theory. I believe these levels of integration are rooted in the physiological as well for Sherrington, although Sir John Eccles has argued that Sherrington subscribed to Cartesian dualism.

For Descartes, a human is a union between two distinct sorts of substances: mind and body. The body is part of the mechanical universe and its workings are explicable solely in terms of causo-physical laws. The mind on the other hand is a pure thinking substance, whose main attribute is thought. The body's main attribute is extension in space. The mind resided inside the body and permeated each part, thus making it sensitive to the body, and also enabling its controlling influence over the clock-work mechanism. The well-known problem with Cartesian dualism is its failure to explain how mind and body interact. If the mind is nonspatial how can it exist anywhere, let alone in one particular body and not in another? And if sense can't be made of its location within locatable bodies, how can it influence a body? In short, if the two kinds of substances are absolutely distinct and share nothing in common, neither can affect the other. Since humans evidently can be described in terms

of mental and physical qualities, and since each can influence the state of the other, Cartesianism fails to explain the phenomenon of human nature. On the other hand, dualism has theological appeal since it enables the possibility of disembodied existence -- if mind is ontologically distinct from body, it can exist without it -- which is necessary if such theological notions as 'heaven', 'purgatory', 'transmigration of souls', 'reincarnation', etc. are to have any meaning or use. I believe theological interests motivated Eccles to interpret Sherrington as being a dualist, although I will try to show that such a view is not the only interpretation, and is not the one supported by the facts.

In a style which tells of his poetic imagination, Sherrington wrote:

> "We turn to behaviour of a differnet kind, some say even of a different category of act. The field of the psyche is entered. An old adage has it that to the trodden worm its own trodden self is the world's greater half. That anthropomorphic worm may typify ourselves to us; the 'self' of each of us goes far to epitomize the integration we are now to look at. We can retain the scheme of spatial nervous arrangement we used before, this time, however, not mutilating the central organ, but keeping the animal -- the human animal if you will -- intact. The receptors at the starting-points of the nerve-threads we find now to be, by conspiracy with the psyche in the central organ, sense-organs. The full panel of the 'five-senses' is in session, and by further collaboration with the psyche, a world of subject and object for the individual is in being. The individual has attained a psychical existence. Phases and moods of mental life accrue. Each waking day is a stage dominated for good or ill, in comedy, farce or tragedy, by a *dramatis persona*, the 'self'. The continuity of its presence in time, seems hardly broken by sleep, its inalienable 'interiority' in (sensual) space, its consistency of view-point, the privacy of its experience, combine to give it status as a unique existence. Although multiple aspects characterize it it has self-cohesion. It regards itself as one, others treat it as one. It

> is addressed as one, by the name to which it answers. The Law and the State schedule it as one. It and they identify it with a body which is considered by it and them to belong to it integrally. In short, unchallenged and unargued conviction assumes it to be one. The logic of grammar endorses this by a pronoun in the singular. All its diversity is merged in oneness." (IANS, p. xiii-xiv)

The solidarity of one's soul, one's 'self-cohesion', is not a matter of grammatical convention, nor of legal fiat; psychical solidarity *too* has its basis in the physiological. Sherrington examined the basis of dependence in 1890, and his findings are also reported in Lecture X where he discussed the relation between the physiology of vision, and the psychology of visual experience. However, the core of his report draws attention to an apparent indifference of perceptual experience to its physiological basis. For example, the psyche treats visual information received from two monocular sources as one binocular field. The psyche "unconsciously takes for granted that its seeing is done by a cyclopean eye having a center of rotation...midvertical of the forehead at the level of the root of the nose." (IANS, p. xiv) Moreover, if each eye is illuminated equally and separately, the subject does not perceive double the brightness, rather, there is no perceptible difference between the effects of this condition and the effects of being informed by either source with both eyes. Again, different colours perceived separately by each eye are combined into one colour, On these findings

Sherrington concluded:

> "Conjunction in time *without* necessarily cerebral conjunction in space is thus an element in the unification of the mind. Simultaneity will of itself make a mental unity. It is somewhat as if two persons of similar makeup could pool their separate psychical experiences into one." (IANS, p. xv)

This is an important conclusion that is perhaps the source, or one source, of confusion over Sherrington's conviction on the metaphysics of mind and body. The disparity between physiological information that originates from spatially distinct sources is negated in regard to spatiality, quantity and quality, when that information is experienced by the percipient being. How the brain overcomes the disparity is a mystery today, but Sherrington knew why this capacity is a necessary condition of perceptual experience.

> "The well-known outline figures, often called equivocal-figures, with which while we gaze at what depicts for instance an overhanging eave the interpretation suddenly changes to a set of ascending steps, have the character of giving always wholly either one thing or wholly the other. The meaning is never at the same time partly this and partly that. Doubtless because to be so would be to have no meaning. Psychical integration is immensely influenced by meaning." (IANS p. xv)

The brain's capacity to ignore 99% of the information it receives renders order to perception and meaningful experience possible. (Guyton, 609) It is important to stress that the brain negates the *disparity* and not all of the information, and thus percepts *do* evidence spatial properties such as location, size, shape and weight, colour and so forth. Eccles interprets this conclusion, and others, as a basis for metaphysical dualism

between mind and body. At most, however, the above conclusion may embellish our metaphysical knowledge that, for instance, certain mental phenomena, such as mental images, are ordered temporally and occur sooner or later than each other and sooner or later than events within full public view, but that mental images are nonspatial; although they evidence some semblance of colour, shape and size, mental images cannot be subject to physical changes, at least directly, as genuinely spatial objects can. The point is that one can heed such metaphysical truths, realizing that one's mental image of an orange is safe from being quartered and served, without subscribing to imponderable dualism. Perhaps this difference between mental and material phenomena loomed too large in the mind of Eccles as he perused Sherrington, and is to blame for his taking the *disparity* between perceptual experience and its physiological basis to be indicative of a metaphysical *division* between mind and body.

Moreover, such disparity is not unique to psychical solidarity, but also characterizes motor coordination and its physiological basis of integrative action. For example, the limb of a spinal dog treats (i.e. responds to) sensory information received from two distinct receptive sources in the same manner as it treats information that impinges its FCP from a single source, i.e., the information, as electric disturbance, is either threshold and causes a reflex response, or is subthreshold and issues in no reflex response; what matters, as it were, is *whether* certain physiological conditions are met, not *how* they are met. In this respect, the effector organ, the limb, is as indifferent to the disparity between physiological information received from separate sources (of the same modality) as is the mind.

Is the mind a kind of effector organ?

Notwithstanding the physiological basis of psychical integration, the fourth level of integration can be described independently of physiological considerations, as Sherrington has done. Given his analytic/synthetic interest, Sherrington speaks of a fifth or "final integration" that involves the integration of motor solidarity with psychical solidarity to form the solidarity of an individual animal, especially of a human. The final integration effects what may be called "individual solidarity". As with previous levels of integration, *unity* is the central concept in 'final integration', even though that which is integrated shows qualitative diversity:

> "There remains yet another type of integration which claims consideration, although to saddle it upon nerve may perhaps encounter protest. Integration has been traced in two great, and in some respects counterpart, systems of the organism. The physico-chemical (or for short physical) produced a unified machine from what without it would be merely a collection of commensal organs. The psychical, creates from psychical data a percipient, thinking and endeavouring mental individual. Though our exposition kept these two systems and their integrations apart, they are largely complemental and life brings them cooperatively together at innumerable points. Not that the physical is ever anything but physical, or the psychical anything but psychical. *The formal dichotomy of the individual, however, which our description practised for the sake of analysis, results in artefacts such as are not in Nature.* Each such is a quasi-organism which does not resemble ourselves, nor does it, *pace* Descartes, resemble dog or cat. For our purpose the two schematic members of the puppet pair which our method segregated require to be integrated together. Not until this is done can we have before us an approximately complete creature of the kind we are considering. This integration can be thought of as the last and final integration." (IANS, p. xv-xvi; underlining mine)

The above passage draws attention to the qualitative difference between psychical and physical events but stresses their functional unity in organisms found in nature. Sherrington's formal, or linguistic, distinction between mind and body, a distinction he draws frequently and most notably in the Gifford Lectures (*Man on His Nature*, 1940), has been interpreted by previous commentators to mean further that mind and body are metaphysically distinct entities --in the manner of Cartesian dualism. With this assumption, a controversy arose over the intention of the dualistic language in which the theory of final integration is couched. The dispute is whether, in using this language, Sherrington intended poetry or philosophy. Presumably seeking to save Sherrington from philosophic ruin, R. Granit (1966) placed such great importance on Sherrington's "poetic" style in the Gifford Lectures that he was forced to conclude that Sherrington is not the philosopher of the nervous system that Leon Asher had perceived, reportedly in 1931. Ganit upset Sir John Eccles considerably, who sharply critized Ganit and others for failing to carry on Sherrington's alleged crusade against the mindless hordes of materialistic monists. In his most recent book, *Sherrington* (1979, coauthored by W.C. Gibson), Eccles asserts "The general theme of *Man on His Nature* is the dualism of man's nature, body and mind." (151, manuscript). He further claims that Sherrington wanted to reverse the "dominant climate of opinion", which favoured materialistic monism, "when he accepted the invitation to deliver the Gifford Lectures". (Ibid, 152). Moreover, throughout *Sherrington*, Eccles hints that Sherrington's study of the nervous system was the driving force behind his dualistic conception of man, a view which Eccles

himself accepted until reinterpreting his own, and Sherrington's, work in terms of Karl Popper's theory of three worlds: matter, mind, and culture. (*Facing Reality*, Eccles 1970)

Although I agree with Eccles that *Man on His Nature* is an important philosophical work, I do not agree with his view that Sherrington is a proponent of dualism. The undesirable consequences of Eccles' view are manifold. In the first place, Sherrington would represent a reverse in the trend set by scientific inquiry itself toward materialistic monism, beginning almost as soon as dualism received its clearest expression four centuries ago in the work of Descartes. Second, since the trend has grown, and not diminished, in its acceptance among modern scientists and philosophers, Sherrington's theory of integration could not be the historical watershed it is thought to be (v. J.P. Swazey, 1969); the concept would then have enlightened us in physiology only to return us to darker ages in philosophy. Third, materialistic monism lies at the very heart of scientific inquiry; it is the methodological presupposition of all science, and certainly of neurophysiology. If the theory of (final) integration is dualistic it contradicts the very method which, Eccles argues, gives it its strongest support. Fourth, the theory of integration would not only be unsupported by its methodological foundation, it would be imponderable, ultimately. In my view, such consequences would severely diminish Sherrington's landmark stature in the history of ideas. For this reason, I will use the remainder of this appendix to examine some of Eccles' assertions more closely.

I am unable to find a well-argued defence of the view that Sherrington's theory of final integration is dualist; in *Sherrington*, Eccles simply punctuates long quotations used as evidence with snappy

assertions. Even then, however, the evidence cited is often ambiguous, and sometimes the assertions which follow are downright misleading. For example, Eccles quotes the following passage from *Man on His Nature*:

> "An act, from the point of view of *mind*, is a doing with a definite aim. It is integrated from parts and some stand nearer to the aim than others. If I have under the microscope an object which I want a friend to see, my hands move the preparation to bring the tiny object into the center of the microscope's field of view. I am seated; my head is stooped, one of my eyes is looking through the microscope-tube, my other is, by practise, sufficiently detached from the other's act to not be convergent with it, and I am unaware for the time being that it sees anything. The fingers of both of my hands are tentatively pushing the glass-slide slightly this way and that, with the 'purpose' to move the object to the desired spot. I would regard the search with my fingers as the focal part of the integrated act. The sitting, the stooping, the looking, although parts of the act do not dominate it as does the search by the fingers. The rest is background.
>
> It is to the focal act with the fingers that the mind particularly attaches. Not to the fingers as such, still less to the executant mechanism moving them, which is away up the arm. What the mind is concerned with is not the act but the aim. It is more aware of the finger-part than of the rest because the fingers' stands nearest to the aim. The rest is mental fringe, even submental. Finite mind is like a moving focal point which wanders restrictedly within each of us, with an aim. Never at any time pervading much of us. At all times most of us is impenetrable to it. Clearly we must not suppose 'life' and 'mind' are one and the same. The finite life is a phenomenon accessible to sense; the finite mind is not." (*Man on his Nature*, p. 211-212)

Following the above, Eccles writes, "On this basis he (Sherrington) excludes recognizable mind from being an attribute of the countless activities of living organisms." (Sherrington, p. 164) Clearly, however, Eccles' assertion is not supported by the passage just quoted. The passage draws attention to a *difference*

in quality between purposive acts--which are sensible -- and mind that focuses on the aim of such acts -- whose focusing is not within the domain of public scrutiny. In contrast, Eccles' assertion is a claim about the *relation* between an act and the attendant mind. To claim that John is qualitatively different from Sue is one assertion; to say that Sue is not John's wife is quite another. One then expects to find a clear statement from Sherrington in the passages following immediately from the above that would confirm Eccles' assertion. Yet in these passages, Sherrington argues against dualism and in favour of the view that mind is some sort of attribute of body! In his preamble to the passage quoted below, Sherrington distinguishes between lower forms of animals that evidence no recognizable mind, and higher forms that do on the basis of the occurrence of secondary nerve-lines linking simple reflex-arcs together; he then states:

> "It is the coming together of these secondary nerve-lines, their interaction, the setting up of a clearing house dealing with transmitted signals and transmitting them to muscle, in short the organizing an integrated motor act, which contributes the earliest habitation, as Fernel would have said, of mind. The integrated motor act even when lowly contains earnest of mind. Hence comes about the germ of our brain, which we may think carries implication of our mind." (MOHN, p. 213)

Later Sherrington imagines mind as "germinating in the primitive animal as an appurtenance to motoricity ..." (Ibid). Eccles makes no effort to reconcile these statements of Sherrington, which are clearly contrary to a dualistic view of mind and body. An appurtenance, or appendage belonging, to motoricity can hardly be viewed as that which is ontologically distinct from motoricity. Thus mind is not a distinct substance for Sherrington as it was for Descartes.

In my view, Sherrington is trying to express a novel solution to the mind-body problem with his theory of final integration, one that is both sensitive to the unique quality of mental phenomena and yet attempts to explain the occurrence of mind with reference to the organization of matter, i.e., mind is another property that emerges from the integrative action of the nervous system. His clearest endorsement of my interpretation occurs in the very preface to his second edition of *Man on His Nature*.

> "This edition has given opportunity for a certain amount of revision as regards detail. The book stresses the view that man is a product, like so much else, of the play of natural forces acting on the material and under the conditions past and present obtaining on the surface of our planet. If the book succeeds in drawing sympathetic attention to this theme in those who read it, its pages will have handsomely rewarded one of the best hopes of its author."

The passage is dated April 1951, less than a year before Sherrington's death.

In fairness to Eccles, part of the difficulty over interpreting the theory of final integration can be blamed on the language Sherrington must use to express his thought. He wrote the Lectures for an audience of laymen, non-scientists and non-philosophers, and as Eccles observes, Sherrington repeatedly refers to himself as "the man in the street" who wants explanations that are non-technical and use ordinary language. Yet Eccles does not allow his observation due consideration. Ordinary language, the language of the man in the street, is inherently dualistic. Of all theories on the mind-body problem dualism is the easiest to communicate in ordinary language, and indeed, it is even difficult to state the problem without *sounding* like a dualist. The subject-predicate logic of ordinary language

entices the unwary into subscribing to a substance-attribute metaphysics, which is the starting point for dualism. Consider for example the grammatically well-formed sentence:

The table is broad.

"Table" is the subject and "broad" the predicate. A substance-attribute metaphysician would offer further that the table is a substance and broad is one of its attributes. We could examine the table and confirm that it is in fact broad and that broadness, moreover, is one among a collection of attributes belonging to the table. A substance-attribute metaphysics is conveniently parallel to the logical structure of ordinary language, and it can manage the analysis of spatially distinct, or complete, substances such as the locatable objects we encounter in everyday experience. But this language-based metaphysics encounters difficulty if we turn to analyze phenomena wherein time is a main relation around which their properties are organized. Processes with properties which do not exist all at once, but which are disclosed over a period of time, that is, cannot be so analyzed. Ordinary language, however, if taken at face value, will mislead us into believing that processes are like objects. The following sentence is as equally well-formed grammatically as the previous example:

Digestion is painless.

Again, the subject is "digestion" and the predicate is "painless". Painless may be an attribute of digestion, but is digestion a substance in the same way a table was thought to be? Clearly not. Digestion cannot be left in a corner overnight, for example, and found to be

unchanged and intact the following morning. Digestion is a complex of attributes, as is a table, but these are never coinstantiated as they are in a table. Digestion too can be observed, but not all at once; it only emerges as what it is over time.

The point is that if we relied on ordinary language for our metaphysics, we would be misled into believing that processes such as digestion, history, migration etc. are complete entities we can expect to observe directly, and all at once. This may appear to be an exaggeration in the examples chosen, but it does appear to be the kind of mistake in reasoning at work when the nature of mind has been contemplated by some. Body is obviously something more than any local act involving a limb, and likewise we uncritically suppose that attention, perception, imagination and so on are merely local events happening to a greater object we call "mind". Yet the mind is more essentially a process than is the body; bodily processes can cease and there remains something for the anatomist to examine, although little of interest to the physiologist, whereas when mental processes cease there is nothing left for the psychologist to explore. In this way, the mind is more like the process of digestion: when it ceases there may be left traces of its presence, but no object called digestion remains behind.

Finally, if we take dualism seriously, and truly believe that mind and body are two absolutely distinct sorts of whatever, we remove the uncritical basis for the belief that mind is like body, at least to the extent that allows dualism to get off the ground; i.e., to the extent necessary to enable us to speak of a substance, or object, called "mind", that, *like* body, is something more than a collection of dispositions. In short, if one takes ordinary language at face value,

one runs the risk of conceptual bankruptcy.

No doubt a large part of Eccles' difficulty with interpreting Sherrington stems from his own trouble with the concepts of mind and life. In *Facing Reality*, Eccles writes, "Our coming to be is as mysterious as our ceasing to be at death." The mystery is heightened by a dualistic view of man, for whom minds become embodied at conception (?) to form life, and disembodied at death. Sherrington strikes this reader as too modest to claim to have solved such a mystery; in his foreward to the 1947 edition of the *Integrative Action*, he even states that our understanding of the mind-body problem is not improved since Aristotle. Yet I believe Sherrington is responsible for more than he has admitted. His research and his theory of final integration lend scientific credence to the modern view of mind and body and of life, which view is expressed in this succinct passage from the philosopher David Hull:

> "There are not many trends discernible in science, but one of them seems to be the shifting of key scientific concepts from the category of things and substances to the category of properties, especially relational and organizational properties. Life is no more a thing than is time, space, gravity, or magnetism. One might well add *mind* to this list. If life is viewed as an organizational property of certain material systems, then we do not have to worry about where life came from when the first living creatures arose from inanimate substances or where it goes when a living creature dies, any more than we have to concern ourselves with discovering where magnetism comes from when an iron bar becomes magnetized or where it goes when the bar becomes demagnetized. In short, we need not postulate a magnetic heaven as the final resting place for good magnetic fields." (*Philosophy of the Biological Sciences*, 1973, p. 129)

I think that in conceiving the general theory of integration, Sherrington tried to shift the key scientific concepts of 'body' and 'mind' away from the category of 'substances', and therefore in

a direction away from dualism, and toward the category of organizational properties. In the first place, Sherrington was concerned to understand the properties universal to animal behaviour. He wanted to understand behaviour, which is a process; he was not content to know anatomy. Given his physiological interests, 'behaviour', not 'body' becomes the operative concept directing his scientific endeavour. Moreover, behaviour is explained, as we have witenssed, as a phenomenon that necessarily emerges given a certain level of organization in the nervous system. The concept of integrative action explains behaviour as an organizational property of matter.

Secondly, with each progression in levels of neurological organization, the emergent behaviour becomes increasingly more complex and sophisticated until mind and subjective experience emerge. Mind too is conceived as a process which can start the process of motoricity, as exemplified in the anticipatory reactions that are all important to the individual and to its species. Once our categories shift from substances to processes that emerge as organizational properties of "the material", the ultimate process, or integration, is seen to be that which achieves individual solidarity. Yet, in asserting that mind, and life, are organizational properties of matter, one no more denies the qualitative differences between mind and brain, than one denies the obvious difference between, say, first order reflex coordination and the neuronal circuitry that causes it. On pain of sounding trite, if one wants to classify Sherrington on the question of the mind-body relation, it would be best to classify him as being neither a dualist nor a materialist, but as being an *integrationist*.

BIBLIOGRAPHY

1. Burch, R.W., "Functional Explanation and Normality", S.W. Journal of Philosophy, 9, spring 1978, p. 45-53.

2. Cummins, R., "Functional Explanation", Journal of Philosophy, 20, Nov. 1975, p. 741-765.

3. Eccles, J.C., and K. Popper, The Self and Its Brain, Berlin, Springer International, 1977.

4. Eccles, J.C., and W.C. Gibson, Sherrington, His Life and Thought, New York, Springer International, 1979.

5. Eccles, J.C., Brain and Human Behavior, Berlin, Springer-Verlag, 1970.

6. Facing Reality, New York, Springer-Verlag, 1970.

7. The Understanding of the Brain, New York, McGraw-Hill, 1973.

8. Edwards, P., ed., The Encyclopedia of Philosophy, New York, McMillan Publishing Co., 1967, vol. 1, p. 310-318.

9. Essler, W.K., "A Note on Functional Explanation", Erkenntnis, 13, Nov. 1978, p. 371-376.

10. Granit, R., Charles Scott Sherrington: An Appraisal, London, Nelson, 1966.

11. Greenstein, H., "Another Look at Functional Explanation", ITA Humanidades, 10, 1974, p. 255-272.

12. Guyton, A.C., Textbook of Medical Physiology, Philadelphia, Saunders, 1976.

13. Hempel, C.G., Aspects of Scientific Explanation, New York, Free Press, 1965.

14. Hull, D.L., The Philosophy of the Biological Sciences, New Jersey, Prentice-Hall, 1974.

15. Nagel, E., The Structure of Science, New York, Harcourt, Brace and World, 1961.

16. Nagel, E., "Teleology Revised", Journal of Philosophy, May 1977, p. 280-301.

17. Purdon, A.C., "Biological Function", Philosophical Quarterly, 29, Jan. 1979, p. 10-24.

18. Ruse, M., Philosophy of Biology, London, Hutchinson's University Library, 1973.

20. Sherrington, C.S., The Integrative Action of the Nervous System, New Haven, Yale University Press, 1906.

19. Ryle, G., Concept of Mind, London, Hutchinson's University Press, 1949.

21. Sherrington, C.S., Reflex Activity of the Spinal Cord, Oxford, Clarendon Press, 1932.

22. Man on His Nature, Cambridge, The University Press, 1940.

23. Swazey, J.P., Reflexes and Motor Integration; Sherrington's Concept of Integrative Action, Cambridge, Harvard University Press, 1969.