Intern. J. Neuroscience, 1981, Vol. 15, pp. 49–68
0020-7454/81/1502-0049$06.50/0

STRUCTURE AND FUNCTION IN NEUROBIOLOGY: A CONCEPTUAL FRAMEWORK AND THE LOCALIZATION OF FUNCTIONS

PAUL A. M. VAN DONGEN and JOHN H. L. VAN DEN BERCKEN

Department of Pharmacology and Department of Anatomy and Embryology, Catholic University of Nijmegen, P.O.B. 9101, 6500 HE Nijmegen, The Netherlands

(Received April 21, 1981)

A framework of concepts on structure and function of the nervous system is presented, in which we attempt to formulate the intended meanings of some ambiguous statements on structure and function in neurobiology in a set of strictly defined concepts.

Function. The word "function" as generally used has different meanings. Two meanings of the question "What is the function of *B*?" are particularly relevant for this subject: (1) "What does *B* do?" and (2) "Why did *B* evolve?" It is shown that question 1 must be answered before question 2.

Function of brain region S. The question "What does brain region *S* do?" has a meaning identical to the question "What is the relation between inputs and outputs of *S*, at specified conditions of *S*?" The answer to this question is the I/O-function. The I/O-function can be described at different levels, for instance at the molecular or cellular level. When neuroscientists ask "What is the function of brain region *S*?", the intended meaning appears to be either "What is the extra-CNS I/O-function of *S*?," or "Why did *S* evolve?" The "extra-CNS I/O-function of CNS subsystem *S*" is "the I/O-function of *S* in which input and output elements are outside the CNS." In many cases, neuroscientists want to know the "behavioral I/O-function of CNS subsystem *S*," i.e., the "I/O-function of *S* in which input and output elements are outside the organism"; they want to know "what is represented by the inputs of *S* outside the organism, and what are the effects of the outputs of *S* outside the organism." Loosely speaking, when a neuroscientist wants to know "the function of brain region *S*," he wants to know "the meaning of its neural messages, and their behavioral effects." Some examples of generally accepted behavioral I/O-functions of clearly sensory and motor parts of the CNS are presented.

Localization of function. The following conclusions are drawn for "localization of function *F*." (1) "Function *F*" must be specified as "I/O-function *F′*" referring to input and output elements outside the CNS. (2) "Localization of I/O-function *F′*" is a kind of shorthand for "identification of the neural activity representing the states of these input and output elements." (3) If neurons exist whose activity only (or mainly) represents this input ("pontifical cells"), and/or only (or mainly) generates this output ("command neurons") (or, in other words, whose behavioral I/O-function is *F′*), it might be said metaphorically that "function *F* is localized in these neurons." In most cases, the existence of such neurons is still an open question.

Objectives and Point of View

Ambiguities. A great deal of research has been devoted to functional analyses of regions of the CNS and to the "localization of functions." The discussion concerning what it is that parts of the CNS do, is confused because (as will be shown below) key concepts, such as "function," "localization of function," "functional" and "structure", are ambiguous or unnecessarily circularly defined. The need to define these concepts has been felt by many authors: experimental neuroscientists (e.g., Von Holst & Von Saint-Paul, 1963; Stein *et al.*, 1974; Zuelch, 1975, 1976; Stein, 1976), "neuro-engineers" (Gregory, 1961), neuropathologists

---

Please address correspondence to the first author at his present address: Department of Comparative and Physiological Psychology, Catholic University of Nijmegen, P.O.B. 9104, 6500 HE Nijmegen, The Netherlands. A preliminary version of this paper has appeared in P. A. M. van Dongen (1980). "The noradrenergic locus coeruleus. Behavioral effects of intra-cerebral injections, and a survey of its structure, functions and pathology", Thesis, Catholic University of Nijmegen, Nijmegen, The Netherlands.

The authors wish to express their appreciation for the helpful criticism exerted by Dr. A. R. Cools and Dr. P. I. M. Johannesma, and by their colleagues. The idea that neural messages can be best understood by what they represent and generate outside the nervous system, has been formulated earlier by Johannesma, for which he deserves recognition here.

The present address of Dr. J. H. L. van den Bercken is: Department of Special Education, University of Nijmegen.

(Luria, 1966, 1973; Denny-Brown in Rasmussen, 1975), systems analysts (Sagasti, 1970; Ackhoff, 1972), psychologists (Weiskrantz, 1968, 1973; Webster, 1973) and ethologists (Hinde, 1975). Therefore, we will attempt to define precisely the concepts relevant to investigations of the CNS.

Objectives. This article is intended to be a contribution to the discussion on the working of (parts of) the CNS. Our objective is to develop a conceptual framework derived from systems theory, and to apply these concepts to functional investigations of the CNS. The reasons for using systems theory are that the concepts of systems theory are concise and clear, and systems theory is designed to analyze the working of wholes consisting of complexly arranged elements (e.g., systems such as the CNS).

Limitations. The scope of this paper will be limited in several respects. (1) An exclusively causal description of the working of the CNS will be used, because the causal principle is a rule of the game in science (Hospers, 1967, p. 317; see section on "Concepts in System Analysis" for the meaning of "a cause"). (2) The description applies to the "statics" of the CNS's working only, i.e., processes such as learning, functional recovery after lesions, etc., will not yet be taken into consideration. Although such processes are relevant for the working of the CNS, the development of a conceptual framework of the "statics" of the CNS's working must precede the formulation of concepts of the "dynamics" of the CNS's working. (3) We will restrict ourselves to generally accepted knowledge on the working of the CNS. (*Note:* "Generally accepted" and "generally acceptable" mean "accepted by" and "acceptable to the forum of scientists considered to be most qualified to judge the pertinent statement.") The probability of making generally acceptable statements on the CNS's working increases if such statements are based on just a few and generally accepted underlying assumptions.

The Predicament of the Neurosciences

Functional systems and localization of functions. Knowledge of what a part of the CNS does, requires prior knowledge that this part "really" is an entity as far as the aspect investigated is concerned, that is, it must have been investigated and (generally) accepted that it is an entity. (*Note:* What is here called "entity" is often called "functional system," but the latter term will be avoided for reasons mentioned below under "Structural and Functional.") Moreover, it must be made clear what is meant by "what CNS parts do." When one wants to "localize functions," one must be clear on the meaning of "localization of functions." That is, an investigator: (1) must be clear on what he wants to localize, (2) must be sure that its localization is logically possible, and (3) must have criteria for deciding whether one has successfully "localized a function" by experiment or otherwise.

The predicament of neuroscience. Prior knowledge on what a part of the CNS does, is required in order to understand the effects of treatments such as lesions or electrical stimulation (cf. Gregory, 1961; Weiskrantz, 1978; Divac, 1979). However, research is often trapped in a vicious circle: for the right experiments and intelligible results, sufficient prior knowledge is needed, but such prior knowledge can only be gained from the right experiments. In general it is difficult to infer the normal function of a part of the CNS from malfunctioning after its destruction or stimulation (cf. Gregory, 1961; Teuber, 1968; Weiskrantz, 1968; Zuelch, 1975, 1976; Schoenfeld & Hamilton, 1977). Consequently, many seemingly unrelated hypotheses on isolated topics persist in neuroscience, while nobody can formulate practicable, critical experiments. This is "*characteristic of a science that has not yet developed mature theoretical structure and paradigms. . . . (T)he multidisciplinary nature of neuroscience introduces complications with regard to whether neuroscience is potentially an entity, or whether it is intrinsically a collection of sciences identificable only on the basis of a common interest in research on the nervous system*" (Swazey & Worden, 1975). The present conceptual framework for structure and function in the nervous system hopefully contributes to the integration of the different disciplines of neuroscience.

CONCEPTUAL FRAMEWORK

Concepts in Systems Analysis

Introduction. Any conceptual framework is based on a number of concepts that cannot be defined in more primitive terms; for such concepts only a stipulative definition can be given (Hospers, 1967). For the basic concepts "element," "concrete element," "abstract element," "property," "state" and "event" the reader is referred to Bunge (1977b, 1979). The definitions of systems concepts are derived from the definitions presented by DiStefano, Stubberud & Williams (1967), Gilbertson (1968), Klaus (1969), Sagasti (1970) and Ackhoff (1972).

Mathematical function. Since "function" is a key concept in this paper, an explicit definition of "mathematical function" is presented here.

Mathematical function (definition). f is a relation between two sets A and B such that to every member x of A, there is a single element y of B.

Note. This definition is derived from Bunge (1974, p. 15). A mathematical function f is a mapping from A into B: $f: A \to B$; "*the value f takes at x, an element of A, is designated by f(x), in turn an element y of B. That is* $f(x)=y$." (Bunge, 1974.)

Cause (definition). THE cause of event E=the COMPLETE set of states and/or events at a given moment, sufficient for the occurrence of event E at a later time.

Notes. (1) "*According to Mill this is the correct scientific definition of "cause." THE cause (the WHOLE cause) is the set of conditions sufficient to produce the event*" (Hospers, 1967, p. 293). (2) A cause is a state or an event taken from the complete set forming the cause. In this paper, "C causes E" means "C is a cause of E." (For a more elaborated analysis of "a cause" see Bunge, 1977b.) (3) The verb "to cause" has the same meaning as the verb "to produce." Sagasti (1970) and Ackhoff (1972) use the word "produce," which they define as a necessary condition, but for reasons mentioned by Hospers (1967, pp. 279–305), we prefer the above-mentioned definition. (4) A state or event C is said to be a cause of another event E, when the following statements have been confirmed experimentally: (a) C is associated with E, (b) C precedes E, and (c) manipulations increasing C, increase E, while manipulations reducing C, reduce E. (5) If C is a cause of event E, E is called an "effect" of C. (6) "Element F influences element G" means "an event of F causes an event of G."

Representation (paraphrase). The set of states/events B is called a "representation" of A, when A is (normally) a cause of B, AND A can be reconstructed from B.

Notes. (1) Bunge (1974) used "representation" only in its semantic meaning: "constructs" (which are elements of a language) represent "facts." "Facts" are represented as "constructs" via brain processes; consequently, when one wants to deal with the working of the brain, the concept "representation" must be extended to include what is represented by neural activity. (2) Like Bunge's "representation," our "representation" is not symmetrical and not reflexive, but unlike Bunge's concept, our "representation" is transitive: when B represents A, and C represents B, then C represents A. (3) When B represents A, there is a mapping from A into B. The function f, describing the mapping from A into B, is not necessarily a one-to-one correspondence. For instance, while light and sound are represented by activity in the optic tract and the auditory nerve respectively, identical stimuli elicit more or less different responses, and in the absence of light or sound, a maintained activity is present. For that reason, not the strong criterion of one-to-one correspondence, but the weaker (and ill defined) criterion "reconstruction of A from B" is taken. From knowledge of the stimulus, the probability of firing of some neurons can be predicted, and from the firing pattern, one is, *a priori*, able to determine the probability of occurrence of at least a part of the stimulus, and even to reconstruct the stimulus (cf. Johannesma, 1981). (4) State/event A can be represented in various ways: for instance (and this might contribute to a stipulative definition) the optic image of a bird can be represented chemically (on a painting, a photograph or film), magnetically (on a video-tape), electromagnetically (as light), or by a spatio-temporal pattern of action potentials (which is again different in the optic nerve, in the corpus geniculatum laterale and in cortical area 17). These various representations can be transformed ("translated") into one another. (5) In the paraphrase above, the expression "(normally)" is included, because some neural activity represents things that are not present or do not even exist. In fantasies/visions/dreams/hallucinations we may perceive snow when there is no snow, and we may even perceive blue snow (to take an example of Bunge, 1974, p. 85). Yet neural activity representing "snow" (which is normally caused by the presence of snow) is present, and even neural activity that combines "blueness" and "snow." If we know enough of the CNS's workings, we are, *a priori*, able to reconstruct "snow" and "blue snow" from such activity. (*Note.* It can rightly be objected that we have mixed up neural and mental events in this example. This objection can, however, be met by speaking not about "the perception of blue snow," but about "the externally observable correlates of the corresponding neural activity," e.g. a verbal report or a painting.)

System (definition). Two or more elements with a nonempty set of relations between the elements such that the investigator can regard the elements together as an entity, depending on his interest.

Notes. (1) The elements and therefore the system can be concrete or abstract. (2) "System" is a thing chosen by an investigator: the investigator can freely choose a thing and a level ("Levels in Neuroscience") of investigation, but once having made his choice, he is no longer free to choose the boundaries of the system under investigation. For instance in the "block diagram" (Sustare, 1978) of Figure 1, he is allowed to choose either S1, or S2, or S3, etc., as the systems to investigate, because these can be considered as entities, but he may overlook relevant events regarding S1 and S3 together (omitting S2) as a system. (3) In the CNS, no sharp demarcation lines exist between parts that are sensory, integrative and motor. So strictly speaking the demarcation of a sensory system is arbitrary. Taking the visual system as an example: are the regions receiving input from the cortical areas 17 and 18 still visual? When investigators speak about "visual system," they generally do not mention the system's boundaries. (4) The words "elements" and "subsystems" have the same meaning.

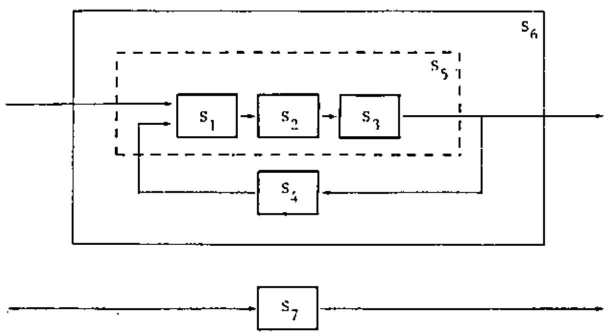

FIGURE 1

Intermezzo: Ontology and systems analysis. In the above-mentioned definition, "systems" can be either concrete or abstract. The logic of the analysis of what systems do, depends on the categories of the elements (are the elements concrete elements, events, "functions," variables, substances, or properties?—cf. Bunge, 1977b, 1979). (Note that one should take care that the elements of the systems and the input and output elements belong to ONE category.) It is a task of epistemology and general systems theory either (1) to develop one conceptual framework which is applicable both to systems consisting of concrete elements (or the concepts representing their working: I/O-functions, cf. Figure 2), and to systems consisting of elements which are states/events (or the concepts representing these states/events: variables, cf. Figure 2), or (2) to show that such a common framework is not feasible. Initially a conceptual framework will be developed for systems consisting of concrete elements (or I/O-functions). At the end of the section "What does *B* do?" the analysis will be extended to systems consisting of states/events, substances and properties.

Structure of system S (definition). The set of relations between the elements of *S*.

Notes. (1) "Structure" both of concrete and abstract systems is abstract. (2) In everyday usage, the word "structure" is ambiguous, meaning either the above-mentioned concept "system" (for instance a part of the CNS) or "structure" as defined above, which is the only meaning of "structure' in this paper. (3) "Structure is identical to "organization" (Klaus, 1969), but only the word "structure" will be used here. (4) By definition, "the structure of CNS subsystem *S*" has a broader meaning than "the anatomical connections of *S*." The structure is not only the connections from *A* to *B*, but also the effects of *A* on *B*.

Input element of system S (definition). An element that does not belong to system *S*, and that influences at least one element of *S*.

Output element of system S (definition). An element that does not belong to system *S*, and that is influenced by at least one element of *S*.

Input of system S (definition). The set of states of the input elements of *S* that influence at least one element of *S*.

Output of system S (definition). The set of the states of the output elements of *S* that are influenced by *S*.

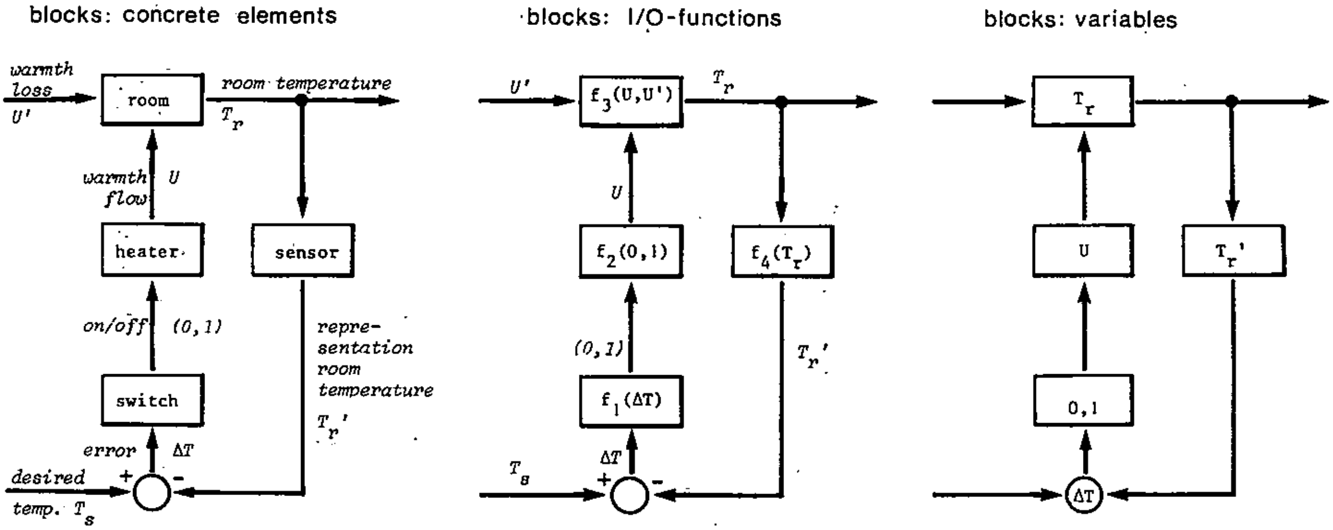

FIGURE 2

Environment of system S (definition). The input elements and output elements of *S*.

Notes. When CNS part *S* is the system under investigation, other parts of the CNS, other organs, or the environment of the organism are the environment of *S*, depending on the level of investigation. On the other hand, when an organism and part of its environment are the system *S* under investigation, the remaining parts of the environment of the animal are the environment of *S*.

Types of Systems

Open system (definition). A system wi.h a nonempty set of relations between elements of the system and of its environment.

Closed system (definition). A system in which all the relations under investigation are relations between elements of the system.

Notes. Open and closed systems should not be confused with open-loop and closed-loop control systems (see below).

Oriented system (definition). An open system with fixed input and output elements.

Nonoriented system (definition). An open system with elements that can be both input and output elements depending on the state of other elements of the environment.

Notes. (1) For a more extensive description of oriented versus nonoriented systems see Lewis (1970); a nonoriented system has been called "circuit" by Lewis. (2) An example of an oriented system is a radio receiver: the aerial is the input element, and the speaker or headphones are the output elements. A radio that can be used both as a receiver and as a transmitter is partially nonoriented: whether the aerial is an input or an output element depends on the user of the radio, who is part of the environment of the radio-system.

Control system (definition). System *S* to keep the state of an element of *S* within preset limits in the presence of changes in the environment of *S*.

Feedback control system (definition). Control system in which a representation of the state of the controlled elements is input of a subsystem which causes a corrective output to hold the controlled variable within preset limits.

Notes. (1) Figure 3A is the conventional diagrammatic representation of a feedback system *S*. In Figure 3B and 3C a representation that is somewhat different from the conventional one is given in order to illustrate more clearly the input and output of the feedback control system (cf., Powers, 1978); sometimes the setpoint is not an input, but an intrinsic state of the control system. (2) Feedback control systems are also called "closed-loop systens." (3) The definition of "control system" can

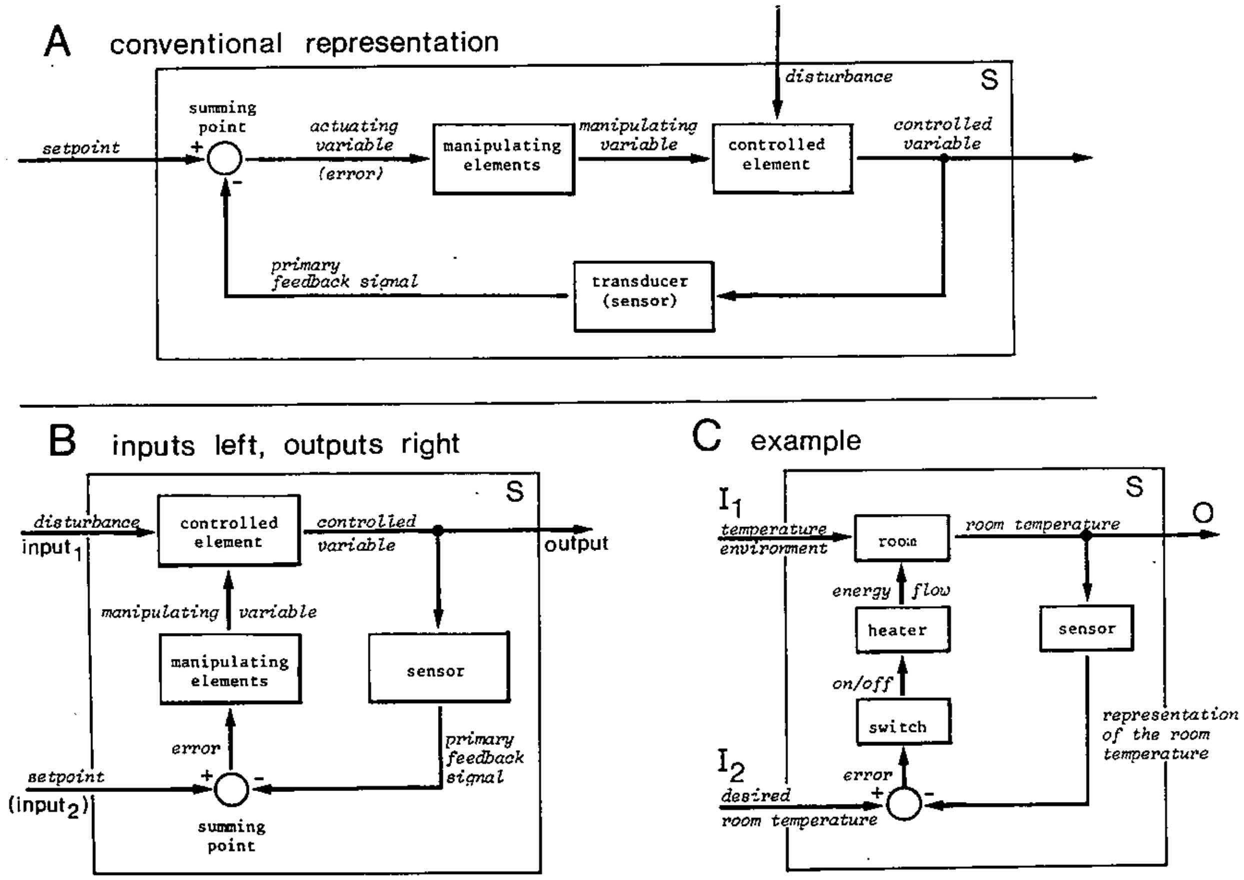

FIGURE 3

be interpreted in a broad sense as encompassing "any system that keeps a state within present limits." It is often convenient to interpret the definition more narrowly as restricted to "systems that have been designed in order to . . ." or to "systems that have been evolved to keep a state within preset limits." The concepts "evolutionary advantage" and "purpose" must be elucidated, before further remarks on control systems can be made (see below sections: "Why did *B* Evolve?", "What is the Purpose of *B*?" and "Control Systems Once Again").

Function

The words "function" and functional" are key concepts in the discussion on the working of the brain. The need to define "function" has been felt by many authors, but an explicit definition rarely has been proposed (cf., Gregory, 1961; Von Holst & Von Saint-Paul, 1963; Luria, 1966, 1973; Denny-Brown, in Rasmussen, 1975; Zuelch, 1975, 1976; Stein, 1976; Schoenfeld & Hamilton, 1977).

Meanings of "function." In neuroscience, the word "function" appears to be used in many different meanings (cf., Luria, 1966, 1973; Hospers, 1967; White, 1968; Klaus, 1969; Sagasti, 1970; Ackhoff, 1972; Woodfield, 1976): (1) Function 1 is the answer to the question "What does *B* do?," where *B* is a system, a state/event, a substance or a property. (2) Function 2 is the answer to the question "Why did *B* evolve?"; in this question, *B* is a part, a property or an activity of an organism, and in some cases the entire organism. (3) Function 3 is the answer to the question "What is the purpose of *B*?," in which *B* is a person/animal, an object or an action. Function 3 is the teleological answer to the question "Why does person/animal *B* do *A*?"

What Does B *Do?*

In a preliminary translation schema, Canfield (1963) suggested: "*A function of B (in S) is to do E, MEANS B does E and that E is done is useful to S*" (see Canfield's final translation schema in the section "Why did B Evolve?"). Consequently, Canfield's concept "function of B" consists of two components: (1) the set of things B does, and (2) those elements from this set that are "useful" to S. In this section the meaning of the question "What does B do?" will be considered. This question will be treated initially for those elements B that are systems consisting of concrete elements (or concepts that describe their working); at the end of this section, it will be extended to those elements B that are states/events, substances or properties.

What does system S *do?* The description of what system S does depends on what kind of system S is. If system S is an open, oriented system, the answer to the question "What does S do?" is the relation between inputs and outputs under specified conditions. Since subsystems of the CNS are indeed open, oriented systems ("Levels in Neuroscience"), the description of what oriented systems do, will be elaborated.

I/O-function (=function 1) *of system* S (definition). The relation between input and output of S, at specified conditions of S.

Notes. (1) This definition is in line with the definition of "function" by Klaus (1969); in other conceptual frameworks (Sagasti, 1970), "the function of system S" is defined as "the set of outputs of S" (redefined in our terms). (2) "Conditions" are the states of elements of S, indicated by variables which are also called "state variables." In organisms, "conditions" are effects either of previous input (e.g., "hunger" after food deprivation) or of internal processes (e.g., circadian rhythms). (3) The I/O-function is an operator describing the mapping from input to output. It can be expressed as a mathematical function: the output of S is a "function of its inputs and conditions": $O=F(I, C)$. (4) The structure of S and the I/O-function of the elements of S determine the I/O-function of S. But a given I/O-function F can be the I/O-function of many different systems S (Klaus, 1969). (5) THE I/O-function of S is the relation between ALL (relevant) inputs and ALL (relevant) outputs of S; an I/O-function is the relations between SOME of the inputs and SOME of the outputs. (6) When S is an animal, an I/O-function is a stimulus–response (S–R) relation. (7) The "function of a tissue" (Luria, 1966) is identical to the "I/O-function of that tissue." (8) Where S is an animal, investigation of I/O-functions (S–R-relations) is what ethologists call "causal analysis" (cf., Hinde 1970, 1974): ethologists search for causation by external factors (inputs) and internal factors (conditions). (9) The "sum of mechanistic operations of system S" (Dawson, 1973) is identical to the "I/O-function of system S"; we prefer the term I/O-function, because "mechanistic operations" sounds mechanical. and our intention is to propose a general description of the working of systems that can also be applied to abstract systems and to information processing systems (such as the CNS), whose workings are primarily understood to be information processing, independent of the physical representation of the information. (10) The I/O-function of S is a representation of the working of S. (11) The I/O-function of a CNS subsystem S is here defined, completely independently of the results of experiments with lesions, electrical stimulation, or chemical manipulations: this paves the way for future comments on "functional recovery," "functional take-over," "redundancy," etc., without circular argumentation.

Mechanism of the working of system S (definition). The causal explanation of the I/O-function of S in terms of the structure of S.

Notes. (1) The mechanism of the working of S is the answer to the question "Why (or how) does S do what it does?" The causal explanation of the working of S consists of the elements of S (with their I/O-function), and the relations between them (the structure of S) (cf. Hempel, 1965). (2) In everyday usage, the word "mechanism" is ambiguous, meaning either "(parts of) a system" (here called "(sub)system" or "element"), "structure" or "the way in which the system works" (the latter is the sole meaning of "mechanism" used in this paper).

Model of system S (definition). Another system, S', whose structure and/or I/O-function is intended to be a representation of the structure and/or I/O-function of S.

Notes. (1) Given a system S with a known I/O-function F and a structure M, many different models S' exist with I/O-function F' and structure M', such that $F'=F$, but this by no means guarantees that $M'=M$ as far as the processes under investigation are concerned (cf. Klaus, 1969; Lewis, 1970). There are no *a priori* reasons (apart perhaps from simplicity) for preferring one model to another. (2) So if an investigator tries to localize a part of the CNS which has an identical I/O-function to that of a sybsystem in his model, he cannot, *a priori*, be sure that he is searching for something that actually exists.

What does state/event, substance or property B *do?* (1) When one wants to know the effects of state/event *B*, one can regard *B* as the input of one or more systems consisting of concrete elements or I/O-functions: the effects of *B* are the output of these systems, and what *B* does is related to the I/O-function of these systems. (Note that state/event *B* can also be considered as the output of another system *S*; in that case the effects of *B* are the outputs of *S*.) (2) Substance *B* can only have effects via its receptors (such is the definition of "receptor"); the concentration of *B* near its receptors is a state: the analysis of effects of substances is thereby subsumed under the analysis of the effects of states/events (see above). (3) Property *B* of an organism can be either: (a) a specified subsystem *S* of that organism, or (b) a specified activity of that organism (i.e., a set of events *E*). What property *B* does is related to the I/O-function of *S*, or to the effects of *E*.

Why Did B *Evolve?*

After an analysis of functional statements in biology, Canfield (1963) concluded: "*A function of B (in S) is to do E, MEANS B does E; and if, CETERIS PARIBUS, E were not done in an S, then the probability of that S surviving or having descendants would be smaller than the probability of that S in which E is done surviving or having descendants.*" The meaning of "*B* does *E*" has been discussed in the preceding section. In the present section comments will be given on the latter part of Canfield's statement: (1) a further analysis of the kind of elements *B* to which such functional statement is restricted, and (2) a further elaboration of "useful," in more general terms than survival and reproduction.

Genes and biological evolution. In this paper (as in biology in general), the question "Why did *B* evolve?" is restricted to those elements *B* which are caused by the presence of different genes. The presence of genes G+ causes the presence of substance g, which would be absent when genes G+ were absent. The effects of the presence of substance g are events and processes which might be, for instance, behavioral patterns or the production of subsystems of the animal. In the question "Why did *B* evolve?" in biology, (1) *B* is restricted to systems, properties and activities of an organism (and in some cases to organisms, see below) whose existence is caused by the existence of G+ genes, and (2) the question "Why did G+ genes evolve?" must be answered by referring to the effects of the presence of G+ genes.

"Usefulness" in evolution. A number of examples have been found of elements *B* in organism *S*, that are caused by genes, but do not contribute to the survival and reproduction of *S*, and even might be disadvantageous for *S*. Arguments have been put forward that the referents ("units") of population genetics and evolution theory are genes rather than organisms, populations or species (Hamilton, 1964; Wilson, 1975: Dawkins, 1976): those *B*'s are favoured in the evolution that cause an increase in the frequency of G+ genes which cause *B*, at the expense of G− genes which cause non-*B* (equivalently, it might be said that G+ genes are favoured).

Evolutionary advantage (definition). The evolutionary advantage of *B* (whose existence is caused by genes G+, and which is a subsystem of an organism) is those subsets of I/O-functions of *B* that caused or still cause an increase in the frequency of G+ genes at the expense of G− genes which cause non-*B*.

Comment. A system *S′* is here dealt with, of which the organism *S* is only a subsystem: *S′* consists of the breeding system of *S* and the biotic and abiotic elements influencing it. The frequency of G+ genes which generate *B* can be regarded as the output of *S′*.

Notes. (1) The I/O-function of subsystem *B* of organism *S* is an intrinsic property of *B* (see above), but which subset of this I/O-function has evolutionary advantage is determined by the larger system *S′*, of which *S* and *B* are subsystems. (2) The evolutionary advantage of state/event, substance or property *B* can be derived from the above-mentioned definition of evolutionary advantage of systems, AND the conclusions mentioned in "What does *B* do?" ("What does state/event, substance or property *B* do?"). (3) In some cases, an organism itself can have evolutionary advantage. Individuals that are either sterile (most male ants), or which are in other ways excluded from reproduction (for instance certain male baboons) can have evolutionary advantage, because their presence and activities increase the probability that other individuals in their breeding system will survive and reproduce which have a high probability of sharing "selfish genes" with the "unselfish" nonreproducing individuals. For this reason, it must not be included in the definition of "evolutionary advantage" that *B* in *S* is "useful" for *S*. (4) The present concept "evolutionary advantage" is related to "adaptation" (cf. Lewontin, 1978; Brandon, 1978), but the latter word is not used in this paper. (5) Of all outputs of *B* only a subset might have evolutionary advantage. For instance, two effects of the beating of the heart are circulation of the blood and the production of heart sounds (Canfield, 1963), the former is evolutionarily advantageous, and the latter is not. (6) In the present definition, a strictly causal formulation is used for evolutionary advantage; all teleology has been removed from this concept (cf. Canfield, 1963; Woodfield, 1976). (7) The concept of "evolutionary advantage" presented here is in line with Luria's (1966) first concept of "function," in which a biological task and the organism's requirements are key concepts. (8) The concept "evolutionary advantage" has the same meaning as the concept "function" (in both its strong and its weak meaning, Hinde, 1975) as used by ethologists (cf. also Hinde, 1970, 1974).

"Why did B *evolve?"* *B* did evolve (and exists), because the G+ genes happened to come into being, and because *B* has evolutionary advantage.

Notes. (1) A part or property of an organism does not necessarily have evolutionary advantage: a part or property without evolutionary advantage may either have been evolved and persist by chance (cf. Kimura, 1968; King & Jukes, 1969), or it may be an effect of the presence of G+ genes which also generate another part or property that indeed has evolutionary advantage. (2) In biology, the question "What is the function of *B*?" often means "What is the evolutionary advantage of *B*?" (cf. Canfield, 1963; Hinde, 1975; Woodfield, 1976): in the answer to the question "What is the function of *B*?" only those effects of the presence of *B* are mentioned that have evolutionary advantage.

What is the Purpose of B*?*

Purpose and intention. The word "purpose" (=function 3) is ambiguous (cf. Hospers, 1967, p. 245), but it means something like "conscious intention." A discussion on mind, consciousness, purpose, reason and goal goes beyond the scope of this paper; the reader is referred to the writings of the following authors: philosophers (Sommerhof, 1950; Nagel, 1961; Canfield, 1963; White, 1968; Woodfield, 1976; Bunge, 1979), systems and information analysts (Sagasti, 1970; Ackhoff, 1972; MacKay, 1972), ethologist (Hinde, 1975), and neuropathologists (Von Cramon, 1978; LeDoux, Wilson, & Gazzaniga, 1979). Man (and animals?) can have intentions/purposes; the purpose of an object or an action depends on their designer, maker, user or performer, and therefore on their purposes. Some comments on the relations between "purpose" and "setpoint" will be made in the section "Control Systems once Again."

"Purpose" in this paper. Relatively little attention has been given to "purpose." This is intentional: our purpose was to develop a conceptual framework for the explanation of the CNS's working limited to causal explanations, that is without invoking teleological explanations. Whether or not such teleological explanations of the CNS's working can be reformulated into causal explanations without loss of their original meaning is an open question.

Structural and Functional

Not only the nouns "structure" and "function," but also the adjectives "structural" and "functional" are ambiguous, so these adjectives must be defined too.

Structural (definition). Of the structure.

Functional (definition). Of the I/O-function.

Notes. (1) In other papers, the word "functional" is used in other meanings (cf. Nagel, 1961; Hempel, 1965; Hospers, 1967; Rosen, 1972; Hinde, 1975; Lahiri, 1977). (2) Localizationistic brain theories are called "structural and functional" or "mechanistic" by Dawson (1973), but there are no compelling *a priori* reasons that a structural and functional analysis of the CNS must result in a localizationistic theory. (3) The adjective "functional" is also used in the meaning "functioning" or "practical"; the concept "functional system" is especially ambiguous, meaning either "functioning system" or "set of functions"; in the present conceptual framework, the concept "functional system" is meaningless.

Control Systems Once Again

"Control system" has earlier (section on Types of Systems) been defined as "a system to keep the value of a variable within preset limits." Now that "evolutionary advantage" and "purpose" have been mentioned, some further remarks on control systems can be made. Such systems are either a part of an organism (for instance the blood pressure control system), or made/designed by organisms (for instance a thermostat-controlled heating system), or they consist of an organism and part of its surroundings (for instance a man–machine system; a concept which is also constructed by an organism). Control systems will be treated in some extension, because: (1) the CNS, the organism and its behavior comprise many control systems, and (2) a clear demarcation of control systems and their input and output can prevent confusion (cf. Powers, 1978).

Controlled state (definition). The state to control which the control system has evolved or been designed.

Controlled element (definition). An element of the control system, whose state is the controlled state.

Controlling system (definition). The control system apart from the controlled element.

Setpoint (definition). The state of the controlled element when between preset limits (goal).

Notes. (1) A stable value of the controlled state either has evolutionary advantage or is the purpose of the designer of the controlling system. (2) Control systems can be either open-loop or closed-loop control systems (DiStefano *et al.*, 1967; Gilbertson, 1968); closed-loop control systems are feedback control systems, which can exercise better control. (3) In some definitions of a feedback control system, the controlled elements are included in the feedback system (Gilbertson, 1968), while in others they are not (Klaus, 1969), and in yet others they may or may not be included (DiStefano *et al.*, 1967). In the present conceptual framework, the controlled element is included by the definition of "system": the relations between the controlled elements and the other elements are of direct concern to the investigator. (4) Control systems (including feedback—closed-loop—control systems) are open, oriented systems (cf. Figures 3 and 4, and see below). (5) The input of a control system is, in all cases, the state of the environment which influences

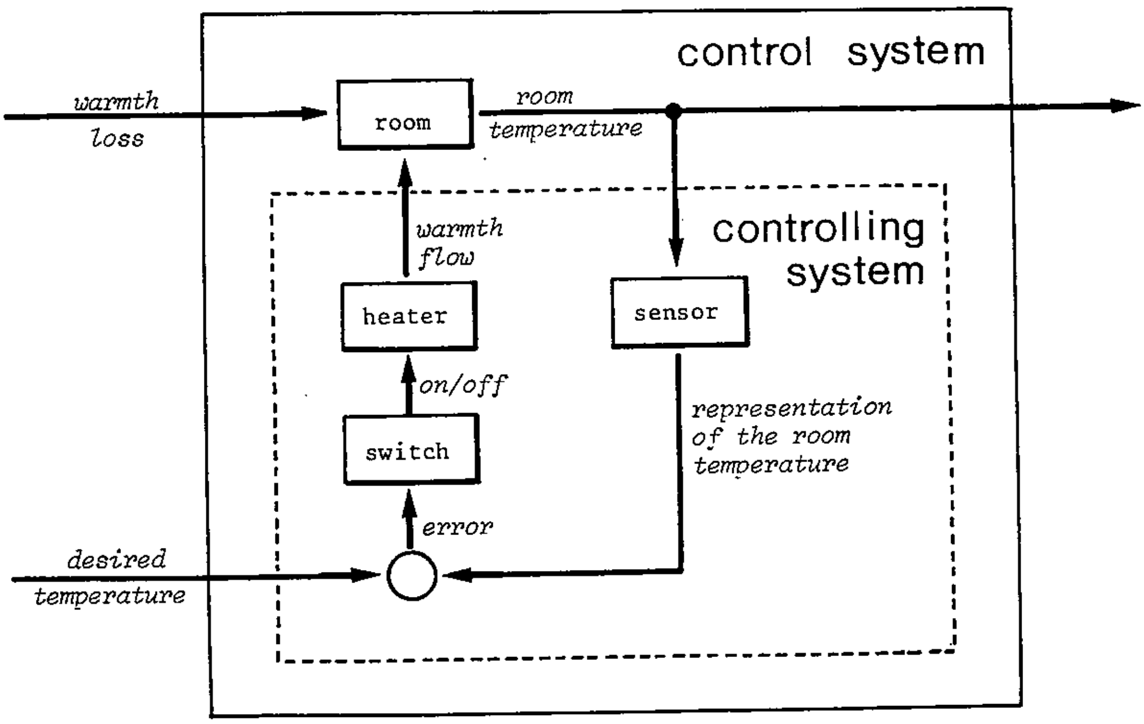

FIGURE 4

the controlled element (generally called "disturbance," cf. DiStefano *et al.*, 1967; Powers, 1978). In some cases, the setpoint is influenced from outside the control system, while in other cases the setpoint is an intrinsic state of the controlling system. (6) The output of a control system is the effect of the states of the controlled element on the environment. When the control system is a part of the organism (for instance the blood pressure regulator), the output of the system is the effect of the blood pressure on other organs than the barosensors (changes in the blood pressure cause changes in all organs). When the control system is an artificial system (e.g., a thermostat-controlled room heating system), the output of the systems affects its user directly or indirectly (the occupant of the room feels the temperature).

Setpoint and purpose. The discussion on whether every purpose can be reduced to setpoints (as defined above), is related to the discussion on the universal validity of the causal principle. Apart from that discussion, attempts have been made to formulate operational criteria to distinguish purposive activities of conscious individuals from goal-directed activities of systems for which there are no compelling reasons for the attribution of consciousness (cf. Sagasti, 1970; Ackhoff, 1972; Woodfield, 1976). The presence of various ways which the system can take to generate its output, is sometimes taken as a criterion of a purposive (or purposeful) system (Sagasti, 1970; Ackhoff, 1972); the latter concept of "function" by Luria (1973) is in line with this concept.

LEVELS IN NEUROSCIENCE

Levels and hierarchies. In the literature a number of brain and/or behavior models can be found, most of which are hierarchical models. The brain is often considered to be a collection of hierarchically ordered feedback control systems (or reflex loops). The arrangement of levels is related to CNS ontogenesis (low is spinal cord, high is neocortex; cf. Bindra, 1976), to psychological theories (TOTE-hierarchies, Miller, Galanter, & Pribram, 1960; low is "intensity control," high is "control of systems concepts," Powers, 1973), or to a comparison of the brain with artificial information processing systems (artificial intelligence, Arbib, 1972). Other hierarchical models which are more or less relevant for this subject refer to communication (low is "vegetative," high is "language," Tavolga, 1970) or behavior (low is "fixed action pattern," high is "behavioral system," Baerends, 1976).

Levels in this paper (definition). In this paper, the following levels will be distinguished in line with Gerard (1958), Bunge (1977a, 1979) and Miller (1978):

i) molecular level: the set of molecules,

ii) cellular level: the set of cells (neurons),

iii) organ level: the set of organs,

iv) organismic level: the set of organisms.

Notes. (1) A "description at level L" means "the elements whose states and events are described belong to L." (2) Still lower and still higher levels could be distinguished, but these are less relevant for this paper. (3) These levels are not hierarchical (Bunge, 1979, p. 14). (4) The reasons for this choice of levels are: (a) this distinction is based on a more general and universally accepted view on the world than other models of brain and behavior, and (b) the levels are unambiguously defined. (5) In this paper, additional levels are used that are "intermediate" to the above-mentioned levels; such "intermediate levels" are not unambiguously defined. Examples are: (a) the subcellular level: the set of subsystems of a cell, consisting of many molecules, and (b) the regional level: the set of CNS subsystems consisting of many cells (e.g., nucleus, region, area).

Application to neuroscience. The conventional use of these levels is "building up the whole world" from low to high levels, e.g., many molecules form a cell, and the actions of these molecules can explain the cell's actions, etc. A somewhat different approach is chosen in this paper: initially a subsystem of the CNS is demarcated (which might be a neuron, a group of neurons, a nucleus, or a cortical area), which will be treated as a "black box." Attention will be paid to inputs and outputs of this subsystem at the various levels: the meaning of "the I/O-function of CNS subsystem S at level L" is thereby defined.

The neuron as a system. A neuron can be regarded as an open, oriented system. Receptors, which are located on the dendrites, the soma and terminals, receive input; the substances which influence these receptors are neurotransmitters and hormones.

Exocytotic ("synaptic") vesicles, located in the terminals and, in some neurons, in the dendrites, influence the neuron's target elements, i.e., the output elements. The concept of a neuron as an open, oriented system can be refined on some points, but in general it seems to be valid. (Note that the neuron is regarded here as an information processing system which handles signals representing something else; incoming substances such as glucose and oxygen, and outgoing substances such as carbon dioxide are not taken into consideration.)

Composition and decomposition of a neuron. A neuron can be divided into subsystems in various ways, for instance (1) in organelles, or (2) in a signal-receiving-integrating-and-spike-generating part (S3 in Figure 5), an action-potential-transporting part S4 in Figure 5), and a signal-transmitting part (S5 in Figure 5). The latter way of subdividing is preferred in this paper for practical reasons: the output of the spike-generating system (i.e., action potentials recorded with a microelectrode near the cell body), and the output of the signal transporting system (i.e., action potentials recorded from fibers in the target region) are directly observable, while the other states/events are not (yet) directly observable or very difficult to observe. (This subdivision is at the subcellular level.)

Input and output at the molecular level. The extracellular fluid around the receptors of a neuron is considered here as its input element; the concentration of molecules (neurotransmitters and hormones) is the input (cf. Figure 5). The receptors on the cell's target elements are the output elements. (The extracellular fluid is not considered as being the output element, because it is also input element of the terminal (cf. Figure 5 and Langer, 1980).) The I/O-function of neuron N2 (in Figure 5) at the molecular level is the relation between the concentration of compounds near the receptors of N2, and the concentration of compounds excreted by N2 near its target elements. Another decomposition of the neuron is often used. The effects of afferent transmitters on the firing rate of neuron N2 is investigated, i.e., the output of S3. Likewise, some investigators study the effects of N2's action potentials on the receptors of S7.

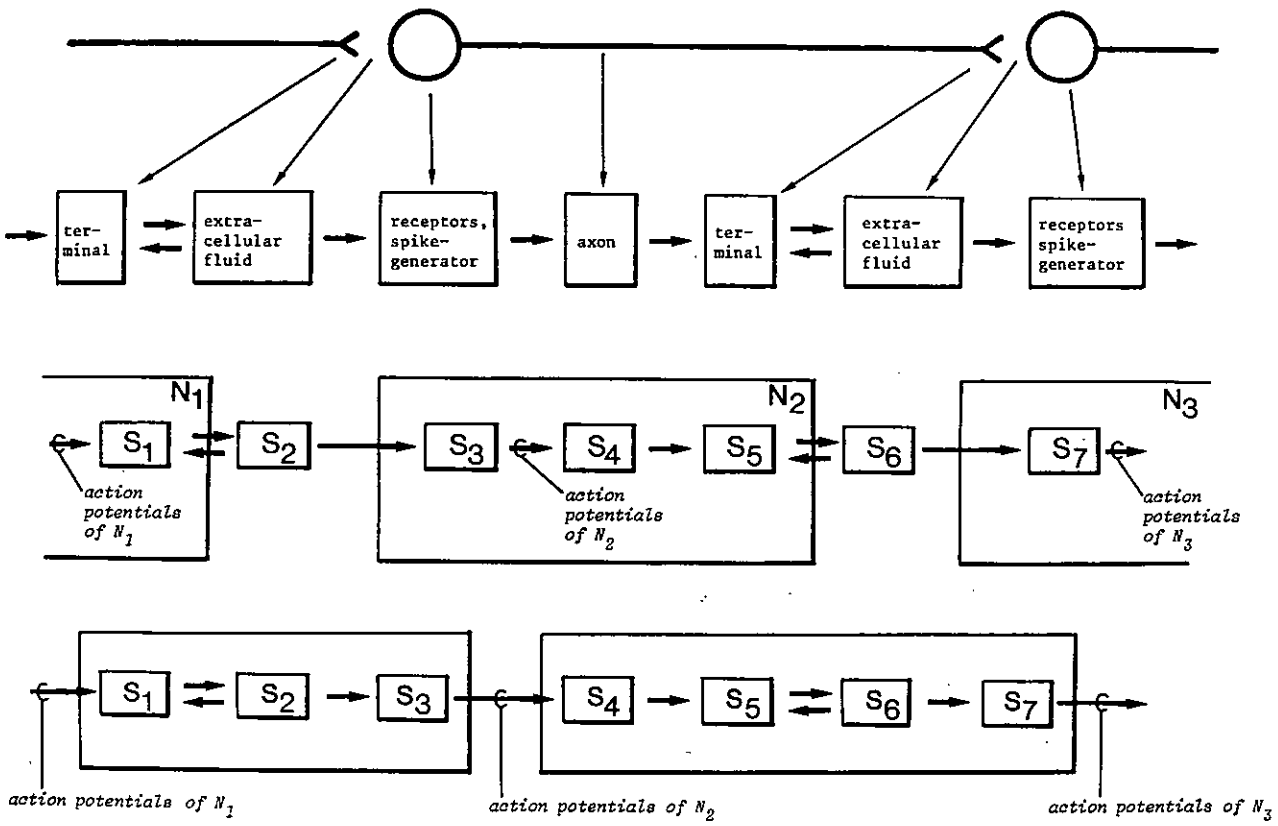

FIGURE 5

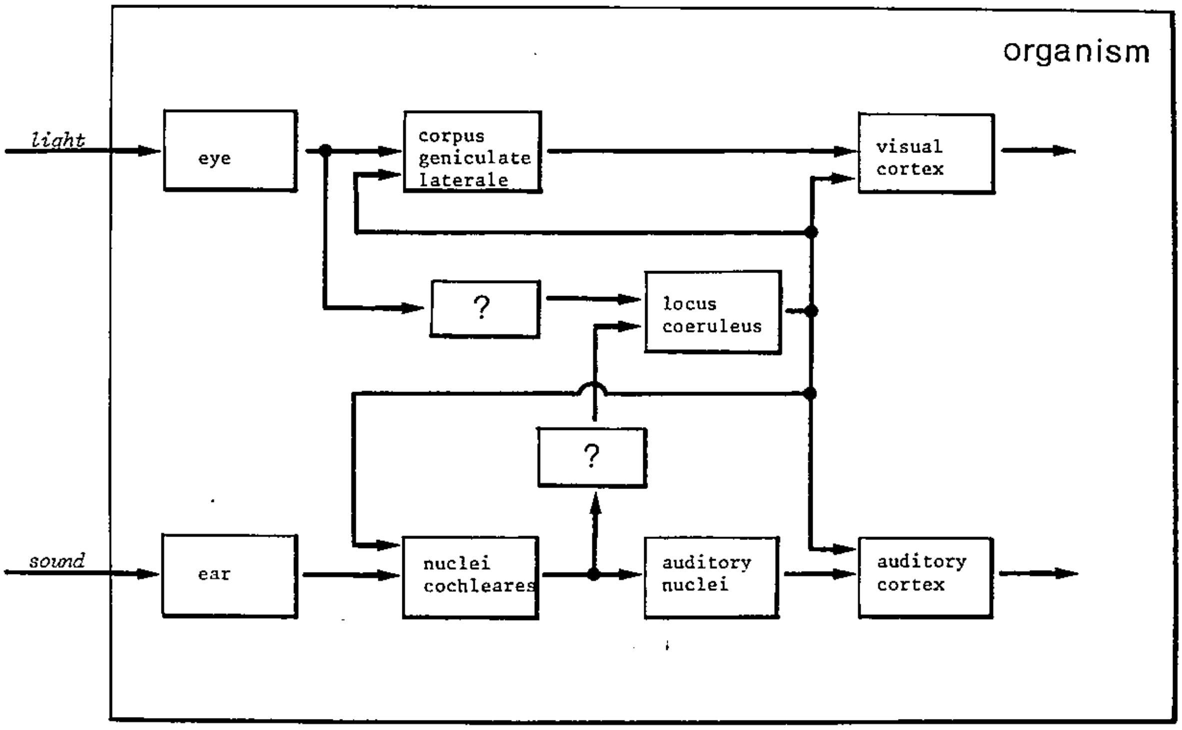

FIGURE 6

Input and output at the cellular level. At the cellular level, the inputs of a neuron are the states/events of afferent neurons (action potentials) or of afferent sensors (often graded potentials). The outputs are the action potentials (firing rate) of its target neurons, or, for instance, contractions of muscle cells. Also in this case, the influence of N1's firing on the firing rate of N2 (in Figure 5), and of the firing rate of N2 on that of N3 are often investigated.

Input and output at the regional level. Sometimes one wants to investigate not only the separate neurons influencing N2, and influenced by it, but one wants to investigate the relations between more neurons simultaneously. Figure 6 gives an example of such investigations at the regional level; such investigation is clearly more complex than at the cellular level: a larger system is now under investigation, and one has to have knowledge of this larger system to deal with it.

Input and output at the organ level. The nervous system is an open, oriented system with sensory and hormonal inputs, and with motor and hormonal outputs. When one wants to study a subsystem of the nervous system (for instance a neuron or a nucleus) at the organ level, the past and/or present states of sensory organs and hormone receptors are the inputs, and the states of muscles, glands and other organs are the outputs. For most subsystems of the nervous system, the input and output at the organ level are indirect.

Input and output at the organismic level. At the organismic level, the interactions between an animal and other animals are relevant. The states/events of other animals are the input via the sensors, and the output is the effects of states/events of the effectors ("response" or "behavior") on other organisms. The input and output of the animal can be described completely at the molecular level (in terms of molecules and light quanta), but not at higher levels, because all organisms consist of molecules, but not all molecules form organisms. When we confine ourselves to the organismic level, only an incomplete description can be given of the I/O-function of a CNS nucleus; for that reason, the "behavioral I/O-function" will be defined in the section "The I/O-Function of a CNS Region," and it will be a key concept in this paper.

"Levels" of signal processing. That the I/O-function of a CNS region can be described completely at the molecular level, but not at higher levels, might seem meagre, because these molecular events together form meaningful stimuli, and that is what matters for the CNS and the animal. Yet at the sensors, the outer world is "degraded" into "molecular events," and that is all the input there is for the CNS. The CNS must "interpret" this sensor-activity into "meaningful stimuli," i.e., it must "reconstruct" an "image of the world." Such a "meaningful image of the world" must again be degraded into cellular events (activation of motor units) to generate behavior. Some theorists have formulated assumptions on how the CNS "makes sense" of the molecular inputs of the sensors, and how the CNS generates "integrated behavior": the various hypothetical stages of information processing have been called "levels" (Miller *et al.*, 1960; Powers, 1973; cf. Tavolga, 1970; Baerends, 1973). We intend, however, to formulate unambiguous concepts to describe experimental data on how the CNS handles sensory input and generates behavior, rather than to formulate specific assumptions on how the CNS might work.

Input, output, levels and experimental brain research. The most simple research situation obtains when the input and output (and thereby the I/O-function) are described at the same level. Many examples, however, can be given of I/O-relations where the input and output elements belong to different levels, especially in experimental brain research. For instance, the input can be at the molecular level (systemic or intracerebral injections of drugs), and the output at the organ level (behavioral effects of these injections). On the other hand, the input can be at the organ level (sensory stimuli), and the output at the molecular level (neurochemical effects of these stimuli) or at the cellular level (effects of these stimuli on a neuron's firing rate). All possible combinations of inputs and outputs at the various levels can be found in experimental brain research.

THE I/O-FUNCTION OF A CNS REGION

Entities in the CNS. When we want to apply the framework developed above to the CNS, a basic question is "Which parts of the CNS are entities (or systems)?", or "On what grounds do we consider parts of the CNS to be entities?" "*Each of a system's elements is connected to every other element, directly or indirectly. Furthermore, no subset of elements is unrelated to any other subset*" (Ackhoff, 1972). Any subsystem of the CNS must consist of connected parts, or in other words, if CNS parts *A* and *B* are connected via *C*, then *A*, *B* and *C* together form the system.

I/O-functions and levels. At a low level, the I/O-function of system *S* can be formulated independently of the larger system *S'* of which *S* is a part. For instance, in the nervous system, the I/O-functions of all cholinergic neurons at the molecular level might be identical: excitatory input increases the ACh concentration at the receptors of their target elements; this applies to the motoneurons of flexors and extensors, the preganglionic cells of sympathetic and parasympathetic neurons, and the septal neurons which project to the hippocampus. The I/O-functions of these neurons differ at a higher level: as yet no general statement can be made about the effects of cholinergic neurons on behavior which is valid for all the above-mentioned cholinergic neurons.

The I/O-function of CNS subsystems and levels. The I/O-function of a CNS subsystem at the cellular level is completely known, when all the afferent neurons, all the efferent fibers, and all the effects of these input elements on these output elements are known. There is a relatively great deal of generally accepted knowledge, for instance, on the anatomy of the hippocampus and the cerebellum, and on how signals are processed in these regions (i.e., knowledge about the structure and the I/O-function of their elements). Yet despite this knowledge at the cellular level, there continue to be many apparently incoherent ideas on what these regions "actually do" (cf. Watson, 1977; Ciba Symposium, 1978). A wealth of knowledge about a CNS subsystem at the cellular level does not necessarily solve the problem what its "function" is. When neuroscientists ask "What is the function of CNS region *S*?" they generally want to know how this region is involved in the actions and reactions of the animal in its environment, i.e., what is called here the "behavioral I/O function" of subsystem *S*.

Extra-CNS I/O-function of CNS subsystem S (definition). The I/O-function of *S* in which the input and output elements are outside the CNS.

Behavioral I/O-function of CNS subsystem S *in organism* S' (definition). The I/O-function of *S* in which the input and output elements are outside the organism.

Notes. (1) The extra-CNS I/O-function is the I/O-function at the organ level PLUS the behavioral I/O-function. In the behavioral I/O-function of organism *S*, the input and output elements are other organisms and nonliving elements influencing *S* and/or influenced by *S*. (2) Some CNS subsystems are (part of) controlling systems, the controlled element of which is part of the body; the working of such systems can be best understood as the I/O-function at the organ level. Also these systems have indirect effects outside the organism. (3) In many cases, neuroscientists want to know the behavioral I/O-function of CNS subsystem *S*. The behavioral I/O-functionof *S* is the "relation between causes and effects outside the organism of activity of neurons of *S*," or in other words, "what is represented and generated by the activity of *S* outside the organism." (4) It is assumed here that a behavioral I/O-function can be formulated for each CNS subsystem, i.e., that each CNS subsystem has (often indirect) effects outside the animal, and that the subsystem's activity is in some way related to the animal's surroundings. The reason for this assumption is that only a system *S* that is part of organism *S'*, can have evolutionary advantage if *S* has effects outside *S'* on the frequency of genes generating *S*. So if a CNS subsystem has evolutionary advantage, it must have effects outside the animal and a behavioral I/O-function. (5) As has been mentioned in the section "Levels in Neuroscience," the behavioral I/O-function can be completely described at the molecular level, but not *a priori*, completely at higher levels. (6) Johannesma has earlier stressed that neural messages can be interpreted by what they represent and generate outside the nervous system; he deserves recognition for this idea. (7) We prefer to speak about "what neural messages represent," rather than "what neural messages mean" (cf. Chung, Raymond & Lettvin, 1970), because the word "meaning" is ambiguous (Hospers, 1967, pp. 11–12); the CNS might, or might not, attribute a "meaning" to a stimulus which is represented in the CNS, independently of whether the stimulus "really" has a "meaning": especially in descriptions of the working of the CNS, "meaning" is a "tricky" word.

Input: what is represented by neural activity? The first stages of information processing in the CNS are obviously related to sensory input. The relationship between properties of the stimulus and the activity of the neuron is described in terms such as "receptive field," "dynamic properties," "trigger feature" or "this-or-that detector." For instance, for retinal ganglion cells, an elaborated classification has been made based on receptive field properties, which appeared to correlate with conduction velocity of the axons, projection areas and form of the cell bodies (Cleland & Levick, 1974a, b; Fukuda & Stone, 1974; Rowe & Stone, 1977, 1979). Such a classification implies that differences have been found in what is represented by action potentials of cells of the different groups. Relatively close to the sensors, cross-correlation between stimulus and neural activity can reveal quantitatively what is represented by neural messages (cochlear nuclei, Grashuis, 1974; Johannesma, 1981; lateral geniculate body, Gielen, 1980); on the basis of such analyses, the responses to new stimuli could often successfully be predicted. It should, however, be noted that both the cochlear nuclei and the corpus geniculate laterale are not only influenced by sensory stimuli, but also by the signals of other nuclei (e.g. the locus coeruleus, cf. Figure 6); so a complete analysis of the activity of these nuclei should also include the effects of activity of other nuclei on them. It is still unknown how complex sensory stimuli are represented further ("higher") in the CNS; speculations have been made based on holographic analogies, mass action and "pontifical cells" ("grandmother cells," cf. Barlow, 1972).

Output: what are the behavioral effects of neural activity? The effects of activity of motoneurons outside the animal are obvious and generally accepted: movements or behavior. The behavioral effects of activity of neurons relatively close to the motoneurons are well studied and generally accepted (e.g., the various spinal reflexes, Lundberg, 1979). As was the case with sensory systems, it is unknown how complex motor/behavioral patterns are generated by the CNS; speculations on "central pattern generators" and "command neurons" have been made (cf. Kupferman & Weiss, 1978; Grillner, 1979).

A theoretical example. Every neural message is a representation of something else, and has a certain (often indirect) effect on behavior. Investigating such representation implies looking-back to what the cause of the activity is. Investigating the effects implies looking-ahead to what is caused by the activity. This is illustrated in Figure 7: a neuron (or a CNS subsystem), N4, receives 3 inputs and generates 3 outputs. N4 is decomposed into a signal-receiving-and-spike-generating part and a signal-transmitting part (cf. "Levels in Neuroscience").

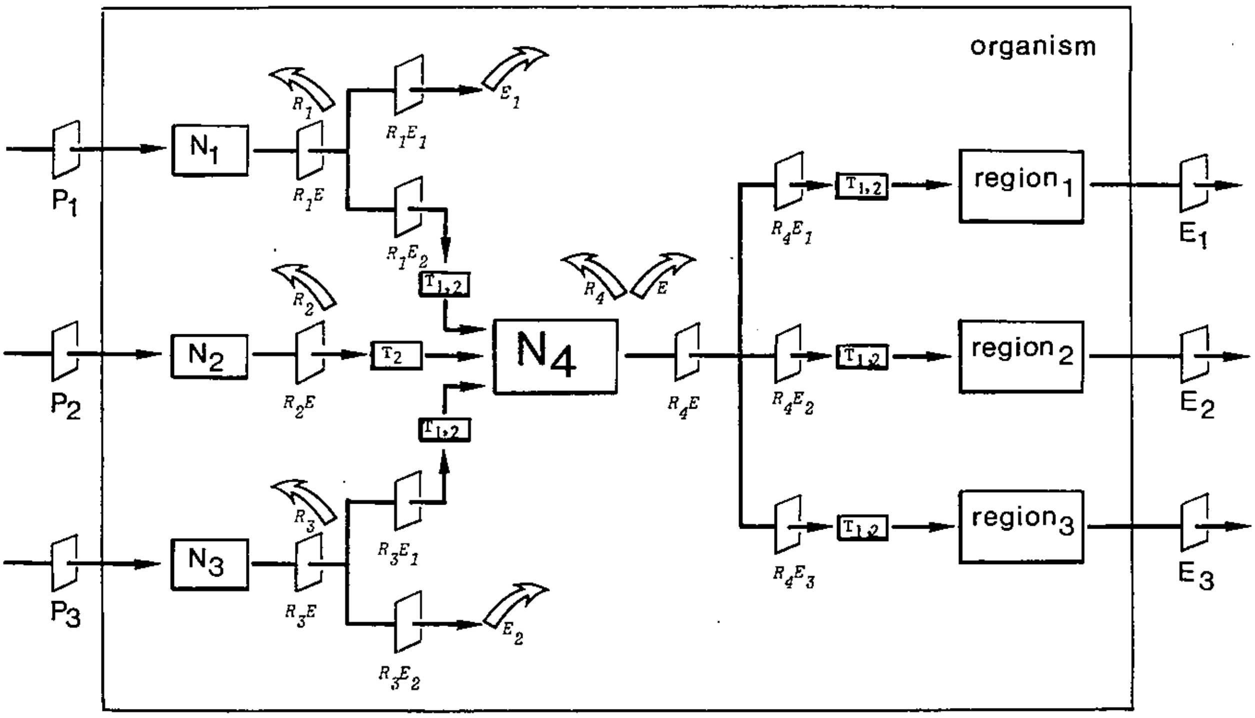

FIGURE 7

It is possible to deduce what the output of N4 represents from what its input (R1, R2, R3) represents and the effects of these inputs on N4 (i.e., R1E2, R2E and R3E1 at the cellular level). The effects at a level L of the activity of N4 (R4E) are deducible from the effects of the inputs at level L (where L is a level higher than the cellular level). The differential effects of activity of N4 (R4E1, R4E2 and R4E3) depend on the different regions to which the output is sent, and on the effects of this output on these regions. In formula, the behavioral I/O-function of N4 is (E1, E2, E3) $=f$(P1, P2, P3). Neurons can be described as operators transforming incoming signals into outgoing signals; these signals can be considered as representations of sensory stimuli, or as generating behavioral effects. The actual nervous system is much more complex, with, for instance, many feedback loops, but in the case of neuronal feedback systems and in the case of artificial ones, one can unambiguously formulate what the various signals represent (cf. Figure 3).

Examples of behavioral I/O-functions. It was attempted some 20 years ago to describe the behavioral causes and effects of neural messages in the frog's tectum (Lettvin, Maturana, McCulloch, & Pitts, 1959); metaphorically, the "bug-detectors" say "there is a prey at that place; catch it" (such activity is a cause, not the cause, of prey catching, cf. Barlow, 1972). More recently, the following hypotheses on the effects of the activity of retinal ganglion cells seem plausible in view of their receptive field properties. For instance, the "on-center sluggish sustained cells" say "the retinal illumination, integrated over a certain area and over a certain time, is high; contract the pupil" (Cleland & Levick, 1974a; Fukuda & Stone, 1974), and some "direction-selective cells" say "the retinal foveal image is moving; make eye movements such that the retinal foveal image becomes stationary" (Oyster, Takahashi, & Collewijn, 1972). (Such activity is a cause, not the cause, of the mentioned effect.) ". . . *the three groups* (of retinal ganglion cells) *fulfill substantially DIFFERENT functional roles in vision*" (Rowe & Stone, 1977). The Ia afferents from the muscle spindles to the alphamotoneurons are a more generally known example: they say "the muscle is too long; contract it." In Van Dongen (1980), a behavioral I/O-function of the locus coeruleus has been proposed, and alternative hypotheses have been formulated in terms of behavioral I/O-functions.

Present knowledge and behavioral I/O-functions. The examples mentioned above of (generally accepted) behavioral I/O-functions of CNS subsystems were restricted to clearly sensory and motor parts of the CNS. Experimental treatments like lesions, electrical stimulation and chemical manipulations have been used to investigate the behavioral I/O-function of other CNS regions. But in a system as complex as the CNS, a complex relationship may be expected between the effects of manipulation of CNS region S, and the behavioral I/O-function of S. In our view, Gregory (1961) was correct with his remark "*To deduce the function of a part from the effect upon the output of removing or stimulating this part we must know at least in general terms how the machine works.*" This idea has been formulated in more general terms: to conclude what a system does, the most important and vaguest point is the amount of prior knowledge of the system (Graupe, 1972). Actually, many apparently incoherent hypotheses exist on what, for instance, the hippocampus, septum, amygdala, hypothalamic nuclei, or central gray do. So it is obvious to conclude that the prior knowledge on such CNS regions is, as yet, too small to make generally acceptable statements on their behavioral I/O-function.

THE LOCALIZATION OF FUNCTIONS

"Function" is an abstract concept, being the representation of the working of a system. Abstract elements as "function" do not exist in space: their location is a meaningless concept. "Functions" are logically not localizable, only concrete elements are. To investigate the cerebral representation of "functions," observable inputs and outputs must be investigated that are directly and unambiguously related to the "function" investigated. When there are lesions in the CNS, a CNS subsystem can do things that it would not have done without that lesion. For instance, in deaf people, understanding of spoken language can be achieved through the visual system (lip-reading). To avoid such unnecessary complications, we will restrict ourselves here to "normal functions" (cf. Luria, 1966, 1973; Rasmussen, 1975). Three examples with different degrees of complexity will be given below.

Examples of Cerebral Localization

Example 1: Respiration. "*Pavlov, when discussing the question of a 'respiratory centre,' was compelled to recognize that 'whereas at the beginning we thought that it was something of the size of a pinhead in the medulla . . . now it has proved to be extremely elusive, climbing up into the brain and down into the spinal cord, and at present nobody can draw its boundaries at all accurately*' " (cited by Luria, 1966). This must be expected for a network as complex as the CNS: if "localization of respiration" means "localization of all the CNS regions involved in, or influenced by, respiration," a great deal of the CNS would have to be included, much more than the original intention of the investigator. "Respiration" is too less specified to state what it is one wants to localize. It is, however, possible to localize, for instance, the blood–CO_2-level controlling system, the input of which is the representation of the actual blood–CO_2-level, and the output of which are effects on the respiratory muscles; and indeed a great deal is already known about its localization. Analogously, the system that controls respiration during speech, or makes a patient exhale when requested to do so by the doctor, can be unambiguously demarcated. To conclude, system S with I/O-function F can be localized, when F is specified in terms of input and output (and often conditions).

Example 2: Vision. Let us suppose that we want to localize "vision," and that we agree that "localization of vision" means "localization of those CNS subsystems whose presence/activity is necessary for vision"; even in that case, it is not clear what we want to localize, because the meaning of "vision" is not clear enough. For instance, after major damage to the occipital cortex, a patient can report seeing nothing, but his eyes can still make movements oriented to visual stimuli, which he denies "seeing" (Poeppel, Held, & Frost, 1973). It is highly probable that he can still make correct eye movements, thanks to intact connections from his eyes to the superior colliculus; the answer to the question "must the superior colliculus be included in the set of CNS regions necessary for vision?" depends on the meaning attributed to "vision." More puzzling still are experiments with split-brain subjects (Sperry, 1974). The left hand of a human split-brain subject can successfully retrieve objects, the written names of which were shown in such a way that they were only represented in the right hemisphere; yet the subject will deny having seen the stimulus and recognizing the objects. "Is the right occipital cortex necessary then for vision?" This question is unclear, because "vision" is not sufficiently specific: it has to be specified in terms of input and output. Some examples of specifying an appropriate question follow. (1) "Which CNS regions are necessary for reactions to visual stimuli?" The answer is simply all regions receiving a projection from the eyes (and in many animals the pineal gland too). (2) "Which CNS regions are necessary for a differential reaction to complex visual stimuli?" The anser is the corpus geniculatum laterale and the cortical areas 17, 18 and 19. (3) "Which CNS regions are necessary for naming visual stimuli?" The answer is probably the angular gyrus (Geschwind, 1979). (4) "Which CNS regions are necessary for naming persons from photos of their faces?" The answer is probably the medial undersides of the occipital and temporal lobes (Geschwind, 1979).

Such questions can now be answered, because the "function" to be localized has been specified in terms of an input (stimulus) and an output (response, reaction).

Example 3: Intermale aggression. "Aggression" is restricted here to intermale aggression, because "aggression" is a group of heterogeneous behaviors, of which intermale aggression is a homogeneous subgroup (Moyer, 1968). This implies that a group of stimuli and movements can be demarcated that are related and that can be regarded as an entity: the intermale aggression system consisting of stimuli and movements. In the CNS there must be a unique set of spatio-temporal patterns of neuronal activity representing the intermale aggressive stimuli, and generating the intermale aggressive movements: a unique, intermale aggression system consisting of CNS activities must therefore exist. This is in any case a set of activities of localizable neurons, but the neurons involved may be involved in other activities too, which would make the concept "intermale aggression system" an incomplete and inappropriate descriptive term of the set of these neurons. In ethology, the input (stimuli), output (behavior) and part of their relations are known, but the structure and the mechanism of the underlying CNS system is not known. Since there is no generally accepted knowledge of how intermale aggressive stimuli are represented in the CNS, and of how complex intermale aggressive behavior is generated by the CNS, we do not know what we are looking for in search of the CNS intermale aggression system. The existence of "intermale aggression pontifical cells" and/or "intermale aggression command neurons" is implicitly assumed in the "localization of intermale aggression."

Conclusions on Cerebral Localization

The "localization of functions." "Localization of function F" means "demarcation of CNS subsystem S which does F"; S can be localized when F is specified as I/O-function F. The questions to be answered experimentally become "Which CNS regions are involved in I/O-function F?" and "Which CNS regions are necessary for I/O-function F?" Note that these are different questions: in a redundant nervous system, CNS region S can be involved in I/O-function F, while being not necessary for F. When S has been demarcated necessary for I/O-function F, the conclusion "F is THE behavioral I/O-function of S" is not justified, but only "F is A behavioral I/O-function of S."

Assumptions in cerebral localization. The story of attempts to demarcate behaviorally defined subsystems in the CNS is a history of failures to formulate generally acceptable theories identifying behaviorally defined subsystems with known CNS subsystems (the I/O-function F of subsystem S of behavioral model M would then be equivalent to the I/O-function F' of CNS subsystem S'); at best, CNS subsystems have been found to be involved in behaviorally defined I/O-functions F. This failure to formulate generally acceptable theories is probably due to the choice of the wrong assumptions: the currently identified behavioral subsystems which are generally accepted, ate probably not equivalents of the generally accepted currently identified CNS subsystems. The "function" one wants to localize might still be localized in, say, 6 "pontifical cells" and/or "command neurons" scattered for instance over various cortical areas (which would give the impression of mass action, while the "function" is strictly ` localized," cf. Barlow, 1972), or it might be not localized/localizable (the corresponding pontifical cells" and/or "command neurons" might be nonexistent).

Conclusions. From the foregoing we draw the following three conclusions: (1) For "localization of function F" "function F" must be specified as I/O-function F' in which input and output elements are outside the CNS. (2) The "localization of this I/O-function F'" is a kind of shorthand for "identification of the neural activity representing this input and generating this output." (3) If neurons exist whose activity only (or mainly) represents this input and/or only (or mainly) generates this output (or, in other words, whose behavioral I/O-function is *F'*), it might be said metaphorically, that 'function *F* is localized in these neurons." The existence of such neurons is in most cases an open question.

REFERENCES

Ackhoff, R. L. Towards a system of systems concepts. In: *Systems Behaviour*, pp. 83–90. J. Beishon & G. Peters (Eds.), London: Open University Press, Harper and Row, 1972.

Arbib, M. A. *The Metaphorical Brain. An Introduction to Cybernetics as Artificial Intelligence and Brain Theory.* New York: John Wiley and Sons, 1972.

Ashby, W. R. *An Introduction to Cybernetics.* London: Chapman and Hall, 1956.

Baerends, G. P. The functional organization of behaviour, *Animal Behaviour*, 1976, **24**, 726–738.

Barlow, H. B. Single units and sensations: A neuron doctrine for perceptual psychology? *Perception*, 1972, **1**, 371–394.

Bindra, D. *A Theory of Intelligent Behavior.* New York: John Wiley and Sons, 1976.

Brandon, R. N. Adaptation and evolutionary theory. *Studies in the History of the Philosophy of Science*, 1978, **9**, 181–206.

Bunge, M. *Treatise on Basic Philosophy. Vol. 2. Semantics I. Sence and reference.* Dordrecht, The Netherlands: D.Reidel Publishing Company, 1974.

Bunge, M. Levels and reduction. *American Journal of Physiology: Regulatory integrative comparative physiology*, 1977a, **2**, 75-82.

Bunge, M. *Treatise on Basic Philosophy. Vol. 3. Ontology I. The furniture of the world.* Dordrecht, The Netherlands: D. Reidel Publishing Company, 1977b.

Bunge, M. *Treatise on Basic Philosophy. Vol. 4. Ontology II. A world of systems.* Dordrecht, The Netherlands: D. Reidel Publishing Company.

Canfield J. Teleological explanations in biology. *British Journal of the Philosophy of Science*, **14**, 285–295.

Chung, S.-H., Raymond, S. A., & Lettvin, J. Y. Multiple meaning in single visual units. *Brain Behavior Evolution*, 1970, **3**, 72–101.

Ciba Foundation, Symposium 58, *Functions of the Septo-hippocampal System.* Amsterdam: Elsevier, 1978.

Cleland, B. G. & Levick, W. R. Brisk and sluggish concentrically organized ganglion cells in the cat's retina. *Journal of Physiology (London)*, 1974a, **240**, 421–456.

Cleland, B. G. & Levick, W. R. Properties of rarely encountered types of ganglion cells in the cat's retina and an overall classification. *Journal of Physiology (London)*, 1974b, **240**, 457–492.

Dawkins, R. *The Selfish Gene.* London: Oxford University Press, 1976.

Dawson, R. G. Recovery of function: Implications for theories of brain function. *Behavioral Biology*, 1973, **8**, 439–460.

DiStefano, J. J., III, Stubberud, A. R., & Williams, I. J. *Theory and Problems of Feedback and Control Systems.* (1967), McGraw-Hill Book Company, New York:

Divac, I. Patterns of subcortico-cortical projections as revealed by somatopetal horseradish peroxidase tracing. *Neuroscience*, 1979, **4**, 455–461.

Gerard, R. W. Concepts and principles of biology. *Behavioral Science*, 1958, **3**, 95–102.

Fukuda, Y. & Stone, J. Retinal distribution and central projections of Y-, X- and W-cells in the cat's retina. *Journal of Neurophysiology*, 1974, **37**, 749–772.

Geschwind, N. Specializations of the human brain. *Scientific American*, 1979, **241**, 158–171.

Gielen, C. C. A. M. *Spatio-temporal and Chromatic Properties of Visual Neurones in the Rhesus Monkey Geniculate Nucleus.* Thesis. Nijmegen, The Netherlands: Catholic University of Nijmegen, 1980.

Gilbertson, G. *Encyclopaedia of Cybernetics.* New York: Manchester University Press, 1968.

Grashuis, J. L. *The Pre-event Stimulus Ensemble. An analysis of the stimulus–response relation for complex stimuli applied to auditory neurons.* Thesis. Nijmegen, The Netherlands: Catholic University of Nijmegen, 1974.

Graupe, D. *Identification of Systems.* New York: Van Nostrand, Reinhold Company, 1972.

Gregory, R. L. The brain as an engineering problem. In: *Current Problems in Animal Behaviour*, pp. 307–330. W. H. Thorpe & O. L. Zangwill (Eds.). London: Cambridge University Press, 1961.

Grillner, S. Interaction between central and peripheral mechanisms in the control of locomotion. *Progress in Brain Research*, 1979, **50**, 227–235.

Hamilton, D. W. The genetical evolution of social behavior. *Journal of Theoretical Biology*, 1964, **7**, 1–52.

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

Hinde, R. A. *Animal Behavior: A Synthesis of Ethology and Comparative Psychology.* 2nd ed. New York: McGraw-Hill, 1970.

Hinde, R. A. *Biological Bases of Human Social Behaviour.* New York: McGraw-Hill, 1974.

Hinde, R. A. The concept of function. In: *Function and Evolution of Behavior*, pp. 3–15. G. P. Baerends, C. Beer & A. Manning (Eds.). Oxford: Clarendon Press.

Hospers, J. *An Introduction to Philosophical Analysis.* Routledge and Keagan Paul, Englewood Cliffs, N.J.: Prentice Hall, 1967.

Johannesma, P. I. M. Neural representation of sensory stimuli and sensory interpretation of neural activity. In: *Proceedings of the Symposium on Neural Communication and Control. Facts and Theories. V Debrecen, Hungary, July 1980.* 1981.

Kimura, M. Evolutionary rate at the molecular level. *Nature*, 1968, **217**, 624–626.

King, J. L. & Jukes, T. H. Non-Darwinian evolution. *Science*, 1968, **164**, 788–798.

Klaus, G. *Woerterbuch der Kybernetik.* Frankfurt am Main: Fischer Bucherei, 1969.

Kupferman, I. & Weiss, K. and comments. The command neuron concept. *Behavioral Brain Science*, 1978, **1**, 3–40.

Lahiri, A. The functional approach in biology. *Bio-Systems*, 1977, **9**, 57–68.

Langer, S. Z. Presynaptic receptors and modulation of neurotransmission: pharmacological implications and therapeutic relevance. *Trends in Neuroscience*, 1980, **3**, 110–112.

LeDoux, J. E., Wilson, D. H., & Gazzaniga, M. S. Beyond commissurotomy: Clues to consciousness. In: *Handbook of Behavioral Neurobiology. Vol. 2. Neuropsychology*, pp. 543–554. M. S. Gazzaniga (Ed.). New York: Plenum Press, 1979.
Lettvin, J. Y., Maturana, H. R., McCulloch, W. S., & Pitts, W. H. What the frog's eye tells the frog's brain. *Proceedings of the Institute of Radio Engineering*, 1959, **47**, 1940–1951.
Lewis, E. R. Neural subsystems: Goals, concepts and tools. In: *The Neurosciences. Second Study Program*, pp. 384–396. F. O. Schmitt (Ed.). New York: Rockefeller University Press, 1970.
Lewontin, R. C. Adaptation. *Scientific American*, 1978, **239, 3**, 156–169.
Lundberg, A. Multisensory control of spinal reflex pathways. *Progress in Brain Research*, 1979, **50**, 11–28.
Luria, A. R. *Higher Cortical Functions in Man*. New York: Basic Books, 1966.
Luria, A. R. *The Working Brain*. Harmondsworth, Middlesex: Penguin Books, 1973.
Mackay, D. M. Formal analysis of communicative processes. In: *Non-verbal Communication*. R. A. Hinde (Ed.), London: Cambridge University Press, 1972.
Miller, G. A., Galanter, E., & Pribram, K. H. *Plans and the Structure of Behavior*. New York: Henry Holt and Company, 1960.
Miller, J. G. *Living Systems*. New York: McGraw-Hill, 1978.
Moyer, K. E. Kinds of aggression and their physiological basis. *Communications in Behavioral Biology*, 1968, **2**, 65–87.
Nagel, E. *The Structure of Science*. London: Routledge and Keagan Paul, 1960.
Oyster, C. W., Takahashi, E., & Collewijn, H. Direction selective retinal cells and control of optikinetic nystagmus in the rabbit. *Vision Research*, 1972, **12**, 183–193.
Poeppel, E., Held, R., & Frost, D. Residual visual function after brain wounds involving the central visual pathways in man. *Nature*, 1973, **243**, 295–296.
Powers, W. T. *Behavior: The Control of Perception*. Chicago: Aldine, 1973.
Powers W. T. Quantitative analysis of purposive systems: Some spadework at the foundations of experimental psychology. *Psychological Review*, 1978, **85**, 417–435.
Rasmussen, T. Localization of normal function. In: *Cerebral Localization*, pp. 293–300. K. J. Zuelch, O. Creutzfeld, & G. C. Galbraith (Eds.). Berlin: Springer Verlag, 1975.
Rosen, R. On the relation between structural and functional descriptions of biological systems. *International Journal of Neuroscience*, 1972, **3**, 107–112.
Rowe, M. H. & Stone, J. Naming of neurones. *Brain Behavior Evolution*, 1977, **14**, 185–216.
Rowe, M. H. & Stone, J. The importance of knowing our own presuppositions. *Brain Behavior Evolution*, 1979, **16**, 65–80.
Sagasti, F. A conceptual and taxonomic framework for the analysis of adaptive behavior. *General Systems*, 1970, **15**, 151–160.
Schoenfeld, T. A. & Hamilton, L. W. Secondary brain changes following lesions: A new paradigm for lesion experimentation. *Physiological Behavior*, 1977, **18**, 951–967.
Sommerhof, G. *Analytical Biology*. Oxford: Oxford University Press, 1950.
Sperry, R. W. (1974). Lateral specializations in surgically separated hemispheres. In: *The Neurosciences. Third Study Program*, pp. 5–19. F. O. Schmitt & F. G. Worden (Eds.). Cambridge, Mass.: MIT Press.
Stein, D. G. La dynamique de la plasticité fonctionelle et morphologique du systeme nerveux central. *Lyon médical*, 1976, **19**, 735–752.
Sustare, B. D. System diagrams. In: *Quantitative Ethology*, pp. 275–311. P. W. Colgan (Ed.). New York: John Wiley and Sons, 1978.
Swazey, J. P. & Worden, F. G. On the nature of research in neuroscience. In: *The Neurosciences: Paths of Discovery*, pp. 569–587. F. G. Worden, J. P. Swazey, & G. Adelman (Eds.). Cambridge, Mass.: MIT Press, 1975.
Tavolga, W. N. Levels of interaction in animal communication. In: *Development and Evolution of Behavior*, pp. 281–302. L. R. Aronson, E. Tobach, J. S. Rosenblatt, & D. S. Lehrman (Eds.). San Francisco: Freeman, 1970.
Teuber, H.-L. Some alterations in behavior after cerebral lesions in man. In: *Evolution of Nervous Central Control from Primitive Organisms to Man*, pp. 170–199. Washington, D.C.: American Association for Advancement of Science, 1959.
Van Dongen, P. A. M. *The noradrenergic locus coeruleus. Behavioral effects of intra-cerebral injections, and a survey of its structure, function and pathology*. Thesis. Nijmegen, The Netherlands. Catholic University of Nijmegen, 1980.
Von Cramon, D. Consciousness and disturbances of consciousness. *Journal of Neurology*, 1978, **219**, 1–13.
Von Holst, E. & Von Saint-Paul, U. On the functional organization of drives. *Animal Behavior*, **11**, 1–20.
Watson, P. J. Nonmotor function of the cerebellum. *Psychological Bulletin*, 1978, **85**, 944–967.
Webster, W. G. Assumptions, conceptualizations, and search for functions of brain. *Physiological Psychology*, 1973, **1**, 346–350.
Weiskrantz, L. Treatments, inferences, and brain function. In: *Analysis of Behavioral Change*. L. Weiskrantz (Ed.). New York: Harper and Row, 1968.
Weiskrantz, L. Problems and progress of physiological psychology. *British Journal of Psychology*, 1973, **64**, 511–520.
White, A. R. *The Philosophy of Action*. London: Oxford University Press, 1973.
Wilson, E. O. *Sociobiology*. Cambridge, Mass.: The Belknap Press of Harvard, 1975.
Woodfield, A. *Teleology*. London: Cambridge University Press, 1976.
Zuelch, K. J. Critical remarks on "Lokalisationslehre." In: *Cerebral Localization*, pp. 3–16. K. J. Zuelch, O. Creutzfeld, & G. C. Galbraith (Eds.). Berlin: Springer-Verlag, 1975.
Zuelch, K. J. A critical appraisal of "Lokalisationslehre in the brain. *Naturwissenschaften*, 1976, **63**, 255–265.