Foundations of Science, special issue on "The Impact of Radical Constructivism on Science", edited by A. Riegler, 2001, vol. 6, no. 1–3: 125–161.

Constructivism, Cognition, and Science – An Investigation of Its Links and Possible Shortcomings

MARKUS F. PESCHL

Dept. for Philosophy of Science, University of Vienna

Sensengasse 8/10, A–1090 Wien, Austria/Europe

e-mail: Franz-Markus.Peschl@univie.ac.at

Tel: +43 1 4277-47624, Fax: +43 1 4277-9476

Abstract

This paper addresses the questions concerning the relationship between *scientific* and *cognitive* processes. The fact that both, science and cognition, aim at acquiring some kind of *knowledge* or *representation* about the “world” is the key for establishing a link between these two domains. It turns out that the *constructivist* framework represents an adequate *epistemological foundation* for this undertaking, as its focus of interest is on the (constructive) relationship between the world and its representation. More specifically, it will be show, how cognitive processes and their primary concern to construct a representation of the environment and to generate functionally fitting behavior can act as the basis for embedding the activities and dynamics of the process of science in them by making use of constructivist concepts, such as functional fitness, structure determinedness, etc.

Cognitive science and *artificial life* provide the conceptual framework of *representational spaces* and their interaction between each other and with the environment enabling us to establish this link between cognitive processes and the development/dynamics of scientific theories. The concepts of activation, synaptic weight, and genetic (representational) spaces are powerful tools which can be used as “explanatory vehicles” for a cognitive foundation of science, more specifically for the “context of discovery” (i.e., the development, construction, and dynamics of scientific theories and paradigms). Representational spaces do not only offer us a better understanding of embedding science in cognition, but also show, how the constructivist framework, both, can act as an adequate epistemological foundation for these processes and can be instantiated by these representational concepts from cognitive science.

The final part of this paper addresses some more fundamental questions concerning the positivistic and constructivist understanding of science and human cognition. Among other things it is asked, whether a purely functionalist and quantitative view of the world aiming almost exclusively at its prediction and control is really satisfying for our intellect (having the goal of achieving a profound understanding of reality).

Keywords: cognitive science, functional fitness, human person, knowledge (representation), (natural) science, representational space

The classical vs. a constructivist perspective of science

What is science[1]? What are its goals? What are its typical means of investigation and construction of knowledge? Why is science playing such an important role in our society? These are only some of the questions which we are posing ourselves, whenever we are facing the phenomenon of science and its implications for our lives and even for our own thinking and perceiving the world.

The notion of *knowledge* about the world seems to be the key to understand the importance of science. As will be developed in the course of this paper the notion that knowledge is ”about” the world (or an aspect thereof) or that knowledge represents/refers to the world is misleading. The (radical) constructivist perspective offers an alternative view: this epistemological relationship of ”aboutness” will be gradually replaced by the notion of a correlation between the world and the internal states of the cognitive system. This does not only apply to a specific kind of knowledge, but to any kind of knowledge.

As our interest is in scientific knowledge we have to be aware of differences to common-sense or everyday knowledge:

(a) Contrary to common-sense knowledge the process of science does not stop its inquiry at determining and establishing (directly) perceivable phenomena;

(b) The *object* of science are the *non (directly) observable* parts of the environment (= “hidden world”) and the goal is the construction of knowledge thereof. This knowledge finds its expression in the form of functional descriptions, predictions, etc. The term “non (directly) observable parts of the environment” refers to entities which cannot be directly perceived by the human sensory system – be it an entity, such as molecular structures or certain ranges of electromagnetic waves, or relations between entities, such as described by theories.

(c) Furthermore, science does not only observe the “non (directly) observable world”, but its goal is to push its investigation further to find out *causes* and *principles* which are responsible for the (dynamics of) the non (directly) observable phenomena.

(d) Finally, science is not only interested in singular facts and their causes, but in *general and universal structures* and *relations*, which are referred to as theories. In many cases these relations and structures are non-material.

Thus, in short science can be characterized as the investigation of causes and as establishing universal knowledge about “hidden mechanisms”. This does not only concern quantitative phenomena, such as described by physics or biology, but also qualitative phenomena, such as the more fundamental questions about what life is or what the qualitative experience of the world is. As will be discussed in section 0, it is only due to modern natural science that the investigation of qualitative phenomena has been reduced to constructing quantitative theories thereof. In the classical view the main goals of science consist in *describing*, *understanding*, and *explaining* a certain phenomenon or a group of phenomena by determining their causes.

Looking more closely at the points above, one can see that things are not as simple as they seem to be. Especially point (b) contains a serious (epistemological) problem: it is by no means clear, how

---

[1] By talking about science this paper refers mainly to the *natural sciences*.

(knowledge about) this “hidden” or “not directly observable world” can be accessed. From science we know very well that we have to make use of *instruments* in order to be able to access this hidden domain of the environment. What does this mean on an epistemological level? An instrument for observation is a kind of *extension* of our own sensory system which has at least three tasks: (i) to penetrate into this hidden part of the environment in order to evoke some desired effect, (ii) to take some measurements there, and (iii) to transform these measurements into a form which can be perceived by the human sensory system. In other words, the environmental phenomenon is not faced in an “epistemologically neutral” way by such an instrument, but with an intention behind it; namely, in order to prove some theory. This implies that this instrument contains already a lot of knowledge stemming from this theory or from “background theories” (cf. e.g., Bechtel 1988, Kosso 1992) – this phenomenon of *theory-ladenness* is one of the main reasons, why such an instrument is not only a device which passively maps the hidden into the perceivable domain, but which *actively transforms* the hidden domain in order to ”prove” a theory and to make it perceivable by our senses. It turns out that the naïve assumption (of realism) about a mapping relationship between the environment and a theory describing it is seriously put into question.

1.1 From realism to constructivism

For the moment, lets have a closer look at the situation in which a scientist and the whole process of science finds itself (see Figure **1**). What we are interested in, is the question concerning the *epistemological relationship* between the environment and its representation in a scientific theory and/or in the cognitive system. Epistemological analysis as well as our own experience suggests that we have to differentiate at least between 4 domains:

H: *hidden domain* having been mentioned above; i.e., this domain is in most cases not directly accessible through our senses and becomes finally the explanandum[2]. This domain has to be assumed, because oftentimes we are confronted with phenomena in our sensible experience (*V*) which cannot be explained satisfactorily by itself;

V: domain of the *“visible” phenomena*: of course this domain is not limited to the phenomena which can be perceived by our visual system, but by any sensory system. Furthermore, it has to be clear that these phenomena are (*emergent*) *effects* of the underlying hidden domain.

R: domain of (primary) *representations*: here the effects of the hidden domain are represented in the cognitive system (i.e., neural representation system).

T: domain of *theories*: this is the realm of science and the representation of a phenomenon in the form of a theory (“explanans”) – be it intentionally (i.e., inside the cognitive system) or externally (e.g., as a linguistic entity, a graphical representation, written paper, book, electronic document, etc.)

---

[2] This hidden domain can be the underlying microscopic or molecular structure of matter, or even the intention of an author, if the object of investigation is a literary text.

Figure 1: The basic situation in which science finds itself when observing the environment: from the hidden domain to the domain of theories.

Through this differentiation we are able to see more clearly the possibilities and differences in epistemological positions concerning the question of the relationship between *H*, *V*, *R*, and *T* in the process of science.

1.1.1 Realistic positions

The positions being more close to realism concern manly the relationship between *H* and *T*. They assume a mapping relationship between the hidden domain and the domain of theories (e.g., scientific realism; Churchland 1979, 1989, 1995). This does not necessarily imply an isomorphic 1:1-mapping, but includes also any kind of n:1-homomorphism; i.e., a class of environmental phenomena/structures is mapped to a single theoretic entity. Such a position has its roots in the (naïve) common-sense assumption that our (common-sense) experience of the world is a mapping of the environment. The supporting argument for such a view is that we are able to (inter-)act successfully with the world by making use of such “realistic representations”. What is left aside is the fact that 1:1 or n:1 representations are not necessarily the only means to generate successful behavior. They are only one possibility in a large pool of other representational strategies. Furthermore, it is by no means clear, how we can determine, when one has found the adequate 1/n:1-relationship, as the question concerning the “epistemological influence” of the cognitive system is more or less avoided. In the position of realism one is inclined to completely detach a scientific theory from its genesis and, hence, from the cognitive processes having lead to it and to focus exclusively on the epistemological relationship between *H* and *T*.

1.1.2 Constructivism I: construction and science

In order to overcome this shortcoming of excluding the genesis of a theory, a constructivist oriented approach takes into account the relationship between the (neural) primary representation *R* and the domain of theories *T*. It seems to be an easily to accept fact that scientific theories are the result of cognitive processes (induction, deduction, abduction, adaptation, etc.). More specifically *T* is the result of an active process of *construction* having its substratum in the neural/cognitive dynamics. In other words, the concept of a more or less passive 1/n:1-mapping between *H* and *T* is replaced by a constructive relationship being realized by construction processes between the levels *R* and *T* (leaving the question of perception and primary representation untouched). Such an account would explain why it is possible that there exists a *plurality of theories* for a single phenomenon which is a well known fact in (the history of) science.

With such a step it is clear that we have to give up the classical idea of a truth relationship between an environmental event and its representation in a scientific theory (e.g., Schmidt 1998). This perspective of ”*scientific constructivism*” forces us to replace the notion of truth with the concepts of “*viability*” or *functional fitness* (e.g., Glasersfeld 1984, 1991, 1995). We shall see that this has crucial implications on what we can and should expect from scientific knowledge.

1.1.3 Constructivism II: cognition and construction

Even as not so radically realist-minded scientists we are inclined to admit that scientific theories *T* are the result of a process of construction which is based on some kind of inductive mechanism extracting regularities out of the visible domain *V* in order to make claims for the hidden domain *H*. (Radical) constructivism pushes these claims having been made in section 1.1.2 even further: whereas “scientific constructivism” maintains the constructive character of scientific theories, “*cognitive constructivism*” goes further by posing constructive processes on the more fundamental level of cognition and representation. I.e., already our representation and experiences of the world are the result of constructive processes taking place in our sensory and nervous system (e.g., Peschl 1994a, 1994b,1997). In other words, body structures and the architecture of the nervous system become theory-laden instruments (in the above sense) – they can be interpreted as evolutionarily developed theories which are embodied in the neural/body substratum.

Construction on the level of cognition has an interesting epistemological implication: namely, that the domain of “visible effects” *V* would suddenly “disappear”. I.e., what we are experiencing as a “visible effect” of the hidden domain *H* is the result of a construction process being realized by the sensory and nervous system: the hidden phenomena are directly interacting with the sensory surface and through that perturbate (e.g., Maturana 1980) the neural dynamics. Thus, it is only through the action of our nervous system that we have the experience of objects, entities, colors, (meaningful) sounds, etc. On a physical level we are only confronted with a more or less structured flow of energy perturbating the dynamics of the nervous system. Whatever the neural dynamics generates out of this flow of energy (be it representations, behaviors, etc.) is the responsibility of the structure of the nervous system. Thus, the representation of the world (either as conscious experience or as a scientific theory) becomes *system-relative*. Furthermore, representation is not so much ”about” a certain state of the environment, but there is a relationship of correlation between these two domains; i.e., knowledge is correlated to the world in a system-relative manner. Whereas in the realist position the role of *H* is to determine *R* and *T*, it only *modulates* the representational domains in the constructivist perspective in order to satisfy this correlational relationship.

1.2 “Cognitive constructivism”: an explanatory vehicle for scientific processes?

If our focus of interest in understanding science is on the dynamics of scientific knowledge and on its development and construction, it seems that the view (of “cognitive constructivism”; section 1.1.3) can be used as an *explanatory vehicle* for this undertaking. It is clear that it is not only an isolated cognitive system which is making science – such a system is embedded in a whole cultural, linguistic, and social environment. However, the goal of this paper is *not* so much to focus on these social or cultural factors, but rather on the epistemological core: namely, on the cognitive system itself and its ability to *construct* (common-sense, linguistic, scientific, etc.) *representations* and to use the knowledge/theories about these abilities as a foundation for our investigation of the process of science. Thus, what we are searching for is a *cognitive foundation for science*. Despite the fact that classical approaches in philosophy of science, such as logical positivism, Poppers philosophy of science, etc., do not include or even explicitly exclude the activity of cognitive systems, it simply seems to be more interesting to focus on the “context of discovery” rather than on the “context of justification”. Taking more seriously this connection between cognitive and scientific processes is a recent development in philosophy of science (e.g., Brakel 1994; Churchland 1989, 1991, 1995; Giere 1992, 1994). Maturana (e.g., 1998), Schmid (1998) and others discuss this problem in the context of a constructivist epistemology.

So, what we are looking for is (i) an epistemological basis which offers an adequate framework for such an integration/reduction. The *constructivist* perspective having been sketched above could provide such a framework – its details in the context of our questions will be discussed in sections 1.3f. (ii) Furthermore, a set of theories is necessary which provides a *tool* for better explaining and understanding the process of science as being rooted in cognitive activity. Concepts recently developed in the fields of *cognitive science* and *artificial life* could provide such tools. As will be shown in sections 1.5f, I do *not* refer to the classical concepts of symbol manipulation or propositional approaches in cognitive science (e.g., Newell et al. 1976, 1980; Fodor 1975, 1981; Winston 1992) – they seem to be more closely related to the "context of justification" through their deductive character. Rather, the emphasis will be on the *inductive* side of science; namely, on learning mechanisms, cognitive and evolutionary adaptation processes, construction processes, dynamics of (knowledge) representation, etc. *Neural computation/connectionism* (e.g., Arbib 1995; Churchland et al. 1989, 1990, 1992; Rumelhart et al. 1986; Schwartz 1990; Sejnowski at al. 1988; Varela et al. 1991), methods and concepts from *artificial life* (i.e., genetic algorithms, cellular automata, autonomous systems, etc; e.g., Langton 1989, 1994, 1995; Steels 1995; Mitchell et al. 1994), and the *dynamical systems* approach to cognition (e.g., Port et al. 1995) have brought about a kind of little scientific revolution (in the sense of Kuhn 1970) in the field of cognitive science – they offer a wide spectrum of theoretical as well as technological tools which have a deep impact on an alternative understanding of knowledge (representation) and even on the understanding of the process of science.

It will turn out that these approaches are conceptually as well as epistemologically quite close to and compatible with a constructivist perspective. Furthermore, they help us to better understand that scientific theories are only one form of representation which is embedded in the more general and flexible neural (and evolutionary) representation system of a cognitive system. The focus of interest shifts form the question of an accurate mapping between the environment and theories to the question of constructing and changing theories in conceptual representational spaces (such as the neural weight space or an evolutionary representation space) and how these theories are embedded in the cognitive system's neural and representational dynamics.

The cognitive foundation of science I: searching for epistemological concepts

1.3 Truth, functional fitness, and success in prediction and control

Both science and cognitive systems have rather similar goals: (i) some kind of the *representation* of the world and (ii) to *use* these representations as a strategy in order to better *predict*, manipulate/*control*, cooperate with and make use of the environment. In order to survive (i.e., to keep its homeostatic state going; cf. Maturana 1980) a living cognitive system needs to have some kind of knowledge about its environment so that the organism is able to orient itself in its medium.

Thus, some kind of *representation* of its environment is necessary. The notion of “representation” does not necessarily refer to the classical understanding of representation as some kind of referential mapping of the environmental structures to a representational substratum. Rather, it covers a wide range of *strategies* which help the particular organism to survive by establishing stable and functionally fitting *correlations* between environmental and representational states; these correlations find their expression in includes body structures (such as eyes, motor systems, etc.), neural structures/architectures, and stretch out to scientific theories as a very sophisticated (and hopefully reliable) form of representation. The primary goal of these strategies is to *generate functionally fitting behavior* (e.g., Glasersfeld 1995; Roth 1992, 1994) – in such a view, the aspect of *action* is more important than the aspect of representation (in the classical sense of referential mapping).

Comparing these epistemological concepts of living organisms to scientific theories it turns out that they have (within a constructivist framework) rather similar goals (e.g., Maturana 1991, 1998; Glasersfeld 1995; Foerster 1973; Steier 1991): a scientific theory is not an “objective description” of the environment, but a *strategy* for successfully coping with the environment. (Scientific) *knowledge* is understood as an *instrument* for better *predicting* and *manipulating* the environmental dynamics. The aspect of mapping/representing the environment is only insofar of interest as it helps to achieve a better control over the environmental dynamics. Modern genetics is a good example for this view: of course, a lot of scientific knowledge and models are necessary in order to describe and understand the genetic structure and the biochemical basis – however, the descriptive and mapping character of these theories is only so far of interest as it serves as a *tool* or as a *strategy* for better achieving the goal of manipulating the genetic material or predicting its effects.

Being interested in a more profound understanding of a certain reality is replaced by finding out and describing *functional relationships* and regularities in the environmental dynamics in order to be able to predict and control certain aspects of this reality. In other words, the descriptive, analytic, and explanatory aspect of a scientific theory is reduced to statements which are capturing the environmental dynamics in “if-then”-rules (i.e., if phenomenon *p* finds itself in the state *x*, it is very likely that it will change into state *y*). As the development of natural sciences during the last centuries has shown, such an approach to reality seems to be quite successful in the order of *efficiency*. Yet, it remains a question, if we as humans and our ability to think are really fully satisfied with such a rather functionalist view of the world (see also section 0), and whether it provides the deepest possible understanding of a certain reality… For the moment, let’s stay in the realm of (natural) science, however: what is normally referred to as “objective scientific description or explanation” becomes only a by-product in the constructivist framework; it receives its status as “true knowledge” only because of its superior success in prediction and control. In other words, “truth” is replaced by “functional fitness” and viability (= success in prediction and manipulation; cf. e.g., Glasersfeld 1995; Schmidt 1998, etc.).

1.4 Regularities, construction, control, and functional fitness in science and cognition – epistemological foundations

If we are looking for a cognitive foundation of the process of science, it is necessary to isolate some basic operations and concepts which apply to both domains; constructivism represents an adequate framework for a better understanding of these processes and for establishing compatibility between these two domains:

1.4.1 Extracting patterns and regularities

Both cognitive processes and the process of science are interested in finding *regularities* in the environmental dynamics/structure. In our experience these regularities manifest themselves as patterns in the spatial and/or temporal domain; i.e., before the cognitive system starts to construct knowledge, it discovers that certain phenomena are occurring according to some repeating pattern. These regularities have their origin in the dynamics of the environment, but which pattern is finally extracted and judged as relevant as well as its particular form depends on the structure-determined dynamics of the particular representation system.

These "primary regularities" are extracted by neurally realized *feature detectors* (cf. e.g., Hubel et al. 1968, 1988; Kandel 1991; Essen et al. 1992) on various levels of abstraction. In the scientific domain these feature detectors find their manifestation in a theory-laden view of the world extracting these repeating features of an observed phenomenon which seem to be relevant for the particular theory. In other words, these spatio-temporal patterns are *system (= theory) relative* and represent the basis for the construction of functional relationships.

Construction and induction

As has been shown in section 1.1 both science and cognitive processes are based on the assumption that there exists a "hidden reality" which is – in most cases – not directly accessible by our sensory systems. This domain *H* is responsible for the phenomena and regularities which we are experiencing through our senses. In order to be able to predict and to gain some control over these phenomena it is, thus, necessary to investigate these hidden mechanisms generating the observed (macro-)phenomena. In order to do so the highly inductive neural machinery isolates states, state transitions, and correlations which seem to be relevant for the observed regularities. In an epistemological perspective there are involved several basic steps and levels of construction:

First, there is the construction of the experienced phenomenon itself; as we have seen in section 1.1.3, every organism is confronted only with a more or less structured flow of energy which is detected through the sensory system. It is the task of the sensory system and the neural processing to *construct* some kind of *coherent* and useful (system-relative) *representation* out of these patterns of energy. In the context of science the extraction of a particular phenomenon happens in the course of an experiment when one is isolating (or even manipulating) certain aspects/states of the environment by making use of measuring instruments.

In a second step the cognitive/neural system as well as in science *correlations* between these isolated states, phenomena, objects, etc. are *constructed* – i.e., through feedback mechanisms, learning processes, and statistical methods (in science) regularities in the form of spatio-temporal patterns are discovered in the environmental dynamics. These regularities do not explicitly "lie around" out there in the environment, but an *active* process of *search/discovery* is necessary. Of course, the environmental dynamics follows some kind of regular pattern, but the regularities, which are extracted by the cognitive system/science are primarily *regularities with respect to the* *structure and needs of the representational system/theory*. In other words, the structure of the representation system constructs regularities according to its own regularities which *fit* into the *constraints* of the environmental dynamics.

Finally, some kind of (scientific) "*theory*" or *universal/abstract description* of these regularities is constructed. As we do not have direct access to the hidden domain it is necessary to construct some kind of hypothetical (hidden/abstract) mechanism which governs and "explains" the "visible phenomena". Furthermore and more importantly, the main task of such a hidden mechanism being expressed in a theory consists in being a *tool* for making good predictions and in providing a means for manipulating the environmental dynamics in a desired way.

In any case the constructed knowledge (e.g., theories, models, abstract mechanisms, etc.) has to fit into the dynamics of the environment like a key fits into a lock (cf. the concept of *functional fitness*, Glasersfeld 1984, 1995). The only epistemological criterion (apart from coherence), which has to be fulfilled by everyday as well as scientific theories, is that they are *consistent* with the internal and external *environmental constraints*; i.e., that they fit into the environment. One of the implications is, of course, that there can exist more than one theory which meets this criterion. As long as a certain knowledge or theory can be used in a beneficial way for the survival (in the most general sense) of the organism or a group of organisms (including also scientific communities), it is a functionally fitting or adequate theory about an aspect of the environment. In such a view every (scientific) theory reduces the environment to knowledge about its *functional relationships*.

1.4.2 Prediction, control, and manipulation

What is the *goal* of all these efforts of extracting and constructing knowledge out of the environment? Both science and common sense knowledge finally aim at *making use* of their knowledge for *predicting* and/or manipulating/*controlling* the environmental dynamics. As the main goal (especially of scientific theories, models, etc.) is not so much to "objectively describe" environmental phenomena/dynamics, but to capture their functional relationships, it follows that their primary focus is on the aspect of *action*; i.e., the knowledge has the form of "if-then"-rules (in the most general sense). This form gives us the means for (a) making predictions ("if phenomenon *p* is in state *x*, it is likely that state *y* will follow") and for controlling the dynamics of a phenomenon ("if phenomenon *p* is in state *x* and you want it to reach state *y* you have to apply action *z*"). The basis for these two operations is the knowledge about the hidden domain; more specifically, about its internal states, state transitions, and functional relationships. Thus, both cognitive systems and scientific processes never construct their knowledge just for depicting the environment, but rather for achieving the goal of *externalizing* successful (= functionally fitting) *behavior* for ensuring the organism's survival (in the broadest sense).

1.4.3 Feedback, modulation, and functional fitness in cognition and science

From the perspective of systems theory cognitive and scientific processes follow a structurally similar dynamics: both can be seen as two intertwined *feedback* systems/loops mutually triggering and *modulating* each other. There is an *internal* feedback loop which is realized by the internal representational/neural dynamics and by the dynamics of theory development, respectively. This recurrent loop is responsible for the dynamics of knowledge and theories. The second feedback loop concerns the *external* interaction with the environment; as a result of the internal representational dynamics the cognitive system externalizes behavior and causes changes in the environment which are detected by the sensory system which, in turn, perturbate the representational dynamics. Similarly, a scientific theory “externalizes behavior” by conducting an experiment. In this process the theory or knowledge is tested in the environment. On the input side the results of this experiment are measured and cause a confirmation or the need for change in the representational structure (= theory). Of course, there is a close relationship and interaction between the internal and external feedback loops. The internal dynamics is responsible for generating hypotheses by continuously adapting, constructing, and changing the representational structures (i.e., neural architecture and, thus, the knowledge, theories, etc. being represented in the neural substratum). Furthermore, the internal dynamics is responsible for externalizing this knowledge in the form of behavior, experiments, applying methods, controlling motor devices, etc. On the other hand the internal dynamics is driven in part by the signals, inputs, stimuli, etc. entering the cognitive system or theoretical domain from the environment via the sensory system or via gauges (and their interpretation). As is shown in Peschl (1994, 1997), the environmental input does *not determine* the internal dynamics, but only has *modulatory* influence on the structure-determined representational domain.

The concept of *functional fitness* plays an important role in the context of neural as well as scientific representational processes. From considerations in constructivism there is epistemological, but also empirical (from neuroscience; e.g., feedback connections in the brain; cf. Varela et al. 1991, Roth 1994) evidence that the concept of referential/mapping representations (such as propositions) has to be abandoned. The aspect of generating behavior (or functioning organisms or theories) seems to be the main task of representational structures. Constructivist approaches (e.g., Glasersfeld 1984, 1995; Maturana et al. 1980) refer to this concept as adequate or *viable behavior*. The goal is to manipulate the organism’s internal and external environment in such a way that it is beneficial for the cognitive system’s survival and reproduction by fitting into the external and internal environmental constraints. Constructive and adaptive processes, such as neural plasticity or evolutionary dynamics, have to change the neural architecture in such a way that it is capable of generating viable behavior. Hence, it is not surprising that we do not have real success in trying to find referential representations in any of the representational substrata. A *categorical error* seems to be involved in these investigations: How can we expect to find referential representational structures (such as symbols) in the substratum which is responsible for generating exactly these structures? In other words, the representational mechanisms and substrata being responsible for generating stable referential representations (e.g., neural activities) are confused with their results (e.g., propositions, mental images, etc.).

What are the consequences of such a constructivist view of knowledge and theories in the process of cognition and science? Any kind of knowledge/representation (i) is always *hypothetical*, (ii) it is in a continuous *flow*, and (iii) characterizes the environment only to the extent what it is *not*. Furthermore, (iv) knowledge is always the result of an active process of *construction*, it (v) is *system-relative* and (vi) *functionally fits* into the environmental structures/dynamics rather than depicting it. Thus, it is (viii) a kind of *strategy* for cooperating with the environment. Finally, (viii) there can exist two or more (*competing*) *theories* or strategies which equally well fit into the same environmental constraints. These epistemological implications apply both for knowledge being involved in science and in “every-day-cognition”.

The cognitive foundation of science II: the contribution of cognitive science

1.5 Cognition and (neural) representation

After having discussed the epistemological issues concerning a cognitive foundation of science, this section presents the “conceptual tools” by which such an integration can be achieved. Recently developed approaches from *cognitive science* (CS) and *artificial life* (AL) provide adequate means not only for conceptually implementing such an integration of science and cognition, but also for realizing this integration within the epistemological framework of constructivism. Without wanting to go into technical details, our considerations concern the following approaches and concepts in CS and AL: *neural computation/connectionism* (e.g., Arbib 1995; Churchland et al. 1989, 1990, 1992; Rumelhart et al. 1986; Schwartz 1990; Sejnowski at al. 1988; Varela et al. 1991) and its implications on the understanding on *representation* (e.g., neural representational spaces, distributed representation, learning, etc.), methods and concepts from *artificial life* (e.g., genetic algorithms, autonomous systems; e.g., Langton 1989, 1994, 1995; Steels 1995; Mitchell et al. 1994), and the *dynamical systems* approach to cognition (e.g., Port et al. 1995).

1.5.1 (Neural) Representational spaces as the fundamental concept for an integration of science and cognition

It seems to be an accepted assumption that cognition (and, thus science) is based on the dynamics of neural systems interacting with the environment and with each other. The question of *representation* is at the core of our considerations, as it provides the basis for any process of life, social interaction, as well as for science. Whenever we are talking about (neural) representation mechanisms we have to be aware that any kind of representation is – at least in a systems theory and constructivist perspective – a kind of (non-linear/recurrent) *transformation* of modulatory input-stimuli into motor action/behavior; in other words, representation in the (radical) constructivist sense consists in establishing a stable *correlation* between external and internal patterns aiming at mutually perturabting each other (in a beneficial manner). In abstract terms such a transformation can be described as a recurrent function and/or a set of differential equations. In the neural computation approach these functions are realized as neural computations in a functional description and simulation of the neural architecture, functionality, and dynamics (e.g., activation functions, learning strategies, neural growth [e.g., Cangelosi et al. 1994]).

The most important concept being provided by these newer approaches in CS is the introduction of *representational spaces*: the core idea has its roots in *cybernetics* and systems theory (e.g., Ashby 1964; Wiener 1948; Heiden 1992) and consists in the idea that every physical system (including an organism and a nervous system) can be abstractly described as a system having a set of *states* and *state transitions* according to which the system under investigation behaves. These states constitute a so-called “*state space*” (mathematically speaking it is a high dimensional (discrete) vector space). The state transitions determine *trajectories* through this space which are responsible for the internal and, thus also for the externally observable behavioral dynamics. At each point in time the dynamics of such a system is determined by (a) the current internal state, (b) by the current input, and (c) by the particular state transition which is determined by the structure of the system (in the sense of Maturana´s concept of ”structure-determined”) and chosen by the input.

The idea is to apply the concept of state spaces to (natural and artificial) neural systems in order to describe their dynamics and to achieve a better understanding of their representational capabilities and dynamics (cf. Peschl 1997 for details on the implications for the concept of representation). In other words, a cognitive/neural system is understood as a system consisting of a large number of neural states which create a state space. As neural states are assumed to account for the representational capabilities of an organism, this state space can be interpreted as a *representational space* in which isolated or groups (e.g., whole trajectories) of representational states of the neural system can be found. Depending on the substratum four representational spaces/domains can be distinguished in cognitive systems (see sections 1.5.2–1.5.5) – the overall representational and behavioral dynamics and capabilities of an organism are the result of a *heavy interaction* between the following four levels/domains of representation:

1.5.2 Representation in the activation space

The *activation space* is responsible for holding the *current representational state* of the neural system. It is a *n*-dimensional space, where *n* is the number of neurons/units of the neural system. At any point in time a neuron has a certain activation leading to the description of the representational state of the *whole* neural system as a *n*-dimensional vector of activations (= “pattern” of activations). This vector can be mapped to a single point in the activation space. Thus, such a state/point in activation space can be interpreted as the current representational state of a cognitive system. However, this state does not represent an environmental phenomenon in the traditional sense: first of all many different neural activations contribute to the pattern of activations (cf. “distributed/subsymbolic representation”). Secondly, as most neural systems have a recurrent architecture, the internal representational state is not only determined by the current environmental input (which is supposed to be represented in the traditional view), but also by the *history* of its previous internal states. The current input can only *select* from a set/space of (structure-determined) possible successor states being determined by the neural architecture/weights. Hence, there is no way of guaranteeing a stable referential relationship between the object/phenomenon to be represented and its representation. As an implication of these facts, what is represented in the neural activation space can be characterized as follows: the current state or pattern of activations (= point in activation space) represents a state which *correlates* the currently selected external input and the previous internal states to each other with the goal of generating functionally fitting behavior.

1.5.3 Representation in trajectories in activation space

If one follows these representational/correlational states in activation space over time, a trajectory develops. It is possible to ascribe *representational character* to such *trajectories* (e.g., Horgan et al. 1996; Churchland 1995). In particular, there is the possibility that a trajectory ends up in a cyclic attractor, chaotic attractor, or fixed point attractor (cf. e.g., Hertz et al. 1991) and one can find some representational correlations in these entities.

1.5.4 Representation in the weight space

It is evident that the neural *architecture* plays a crucial role for the dynamics of the spreading patterns of activations (= “current representation”, see 1.5.2). Abstractly speaking the architecture can be captured in a matrix of *synaptic weights* connecting the neurons/units with each other. These weights are responsible for the representational and, thus, behavioral dynamics of the organism – they determine the space of possible state transitions in the activation space. Thus, the whole *representational* and *behavioral potential* (= “the organism’s knowledge”) is *embodied* a particular set of synaptic weights.

What we are referring to as ontogenetic adaptation or *learning* (i.e., some kind of change in the knowledge) is realized as a change of synaptic weights/architecture in neural systems. Similarly as in the activation space, the *m* synaptic weights of a neural system form a state space – in this “weight space” each point refers to a certain state of pattern/set of weights each of them being responsible for a certain set of behavioral/representational dynamics (in the activation space). Thus, the dynamics of adaptation/learning manifests itself as a *trajectory* through weight space. The goal of any learning process is to find such a synaptic configuration (= a certain point in weight space) which determines the dynamics of the activation space in such a way that it is capable of generating functionally fitting behavior. The neural computation/connectionist approach knows a whole set of learning strategies which cannot be dealt with here in detail (most of which are based on the concepts of D.O.Hebb 1949; for a comprehensive overview see Hertz et al. 1991, Arbib 1995). In natural neural systems one can find evidence for this kind of learning processes as well (e.g., long term potentiation, long term depression, etc; e.g. Arbib 1995; Brown et al. 1990, Gazzaniga 1995).

From a representational perspective each *point* in *weight space* represents the *total knowledge* of an organism at a certain point in time and the whole weight space can be characterized as the space of all possibilities (of behavioral/representational/activation dynamics) which can be potentially learned by a particular neural system. The changes in weight space occur according to the feedback of success and failure in the behavioral dynamics being generated by the dynamics of the activation space – thus, there is a relationship of heavy interaction between these two representational spaces.

1.5.5 Representation in the genetic space

The genetic code is responsible for the basic representation of any organism. It does not only determine the structure and equipment of the body, but also the basic developmental strategies and architecture of the nervous system. Of course, the particular “construction” of an organism happens in very complex non-linear processes of *expression* and *development* in heavy interaction with the current environmental dynamics and the already existing body structures (e.g., Edelmann 1988; Berger at al. 1992; Jessel 1991, and many others). From an epistemological perspective the genetic material can be understood as a strategy for generating functionally fitting behavior in the form of a functioning organism (which itself is capable of generating functionally fitting behavior). Via the criterion of success (= *reproduction*) and failure of the generated organism and its behavior the genetic material and the neural representation are linked in a feedback relationship. The genetic drift finding its expression in the evolutionary dynamics can be captured in trajectories in the *genetic space*. Each point in this space represents a particular configuration of an organism giving rise to a particular structure of the weight space itself determining a certain set of representational and behavioral dynamics in activation space. Artificial life provides the tool of various forms of *genetic algorithms* which are capable of simulating these processes (e.g., Holland 1975; Mitchel 1994; Langton 1995, and many others).

Combining genetic algorithms with models from computational neuroscience gives us a powerful set of concepts and instruments for achieving a better understanding of the complex interaction between ontogenetic and phylogenetic processes including "social interaction", autonomous agents (e.g., Steels 1995), or "cultural dynamics" (e.g., Belew 1990, 1992; Miller et al. 1989; Hinton 1987, and many others). As will be shown in the following section they can be applied also for giving us a "cognitively founded view" on science.

1.6 Mapping representational spaces to theory spaces

In order to achieve a better understanding of the *inductive* processes being involved in science the concepts of learning and genetic drift having been discussed above can be used as a powerful tool. Changing the synaptic weights is an inductive process in which "new" knowledge or theories is/are generated on a *trial and error* basis. The resulting hypothetical knowledge acts as a strategy for generating behavior and is applied under the assumption "as if it was true or fitting". Only in the *deductive* process of *externalizing* this knowledge, it becomes evident, whether the theories or knowledge being represented in the current synaptic configuration are/is viable or not. The success or failure of the externalization leads to an internally or externally determined error. This error is an expression of a (n epistemological) mismatch in the functional fitness between the representational and environmental structure. The goal is to reduce this error/mismatch in the following learning steps until the neural architecture embodies sufficiently functionally behavioral strategies. From this perspective learning turns out to be a *search process* in which an error has to be minimized by – abstractly speaking – moving around in the synaptic weight space. A minimal error means a (n epistemologically) stable relationship with the environment in the context of the organism's task to survive.

1.6.1 Representational spaces and theory spaces in science

How are these concepts stemming from cognitive science related to our original question concerning a cognitive foundation of science? Especially the abstract conceptual framework and the rather *dynamic* and *inductive* character of these more recent approaches in cognitive science can shed some light on a domain in (philosophy of) science, which has always been a bit mystified, namely the question of *discovering new knowledge*, the construction of theories, and their dynamics.

As cognitive processes are the basis for any scientific activity, it is not surprising that we can find structurally similar concepts and processes in both domains (see sections 1.4ff). From computational neuroscience and philosophy of science we know well (compare also section 1.5.4f) that the discovery of knowledge in most cases is not so much characterized as discovering something completely "new" in the environment. Rather, it is the construction and discovery of new relationships and strategies with respect to the structure, constraints, already existing knowledge/theories, and needs of the constructing cognitive system. Let's have a look, how these inductive processes in science are embedded in cognitive processes by making use of the concepts having been presented above: first of all, it has to be clear that scientific theories are represented in the same way as any other knowledge in the neural substratum. Therefore, they can be interpreted as certain *states* in a state/representational space.

Mapping what has been said about representational spaces to the question of the representation of a theory in a cognitive/neural system, implies that to represent a certain theory T_i means to be in a certain state in *synaptic weight space*. Why is that so? To assume a certain weight configuration (e.g., in the process of learning) implies a set of behavioral strategies which can be externalized in/to the environment (in certain internal and external contexts). Figure **2** shows this situation for scientific theories on a more abstract level. The (implicitly assumed) scientific paradigm (in the sense of Kuhn (1970) gives rise to a space of possible theories which is embedded in the larger cognitive system's synaptic weight space. Each point in this theory (weight) space instantiates a certain theory T_i. Keep in mind that a single point in synaptic weight space "represents" the total knowledge of an organism at a certain point in time and that a particular scientific theory T_i is only a subset of this total knowledge. Thus, moving around in the synaptic weight space in the course of learning/adaptation has a more or less direct effect on the state of the theory space; and, thus, on the current theory T_i concerning a certain phenomenon. A theory in this sense is always strictly private and has to be communicated and negotiated with the private semantics of other cognitive systems (being interested in the same theory). This is realized by establishing a *consensual domain* (cf. Maturana 1980, 1991). If this process is successful and the resulting theory respects the established "rules" and criteria of scientific knowledge, this theory is ascribed an "objective character".

Figure 2: Embedding scientific theories and their dynamics in the representational space of a cognitive system. Each point in weight space implies the instantiation of a certain theory T_i.

Constructing new theories or changing/adapting already existing theories is based on the same neural processes as any learning process (cf. 1.5.4) which can be characterized as *search process* in a representational state space. In Figure 2 this *cyclic process* of developing theories is depicted in detail: the theory space is embedded in the neural representation space. A certain point in this (high dimensional) space represents a certain theory T_i. Similarly as in the common sense domain, this neural configuration leads to (i) a *prediction* which, in turn, can be (ii) *externalized* in the form of an *experiment*. The experiment can be compared to the behavior of a cognitive system. I.e., some kind of direct or indirect motor action modulates, or *perturbates* the environmental dynamics. In other words, the theory represents knowledge or a *strategy* about *how to penetrate the environment* in such a way that a desired or predicted state/effect is achieved. From this perspective the scientific approach does not really bring about a qualitatively new epistemological procedure for investigating a phenomenon. It becomes evident that scientific processes can be characterized as an extension and refinement of a "purely cognitive" approach to the environment.

The theory also determines the *methods* which are applied to the environment. In the case of cognitive systems the “method” are the motor systems; in the scientific realm these motor systems are extended by more or less complex *tools/instruments* and/or machines which perturbate the environment according to the theory’s rules and instructions. In any case, the resulting “theory’s behavior”, which is externalized in the course of an experiment, fits more or less into the perceived structure of the environment. What does that mean? The level of functional fitness is determined by the success/failure of penetrating and modulating the environment in a certain (desired) way. The *goal* is to *reduce* the *inconsistencies* between the theoretical descriptions, predictions, “behavioral instructions”, and the actual perceived environmental dynamics. Looking at this process the other way around, theories as well as any other (successful) representational structure can be described as results of a process which aims at *establishing consistency* between *environmental constraints*[3] and constraints being determined by the theory, its internal consistency, its goals, and its background framework. As these processes are rooted in neural adaptation processes, the construction of theories can be described as well as an *optimization process searching* for an *adequate transformation mechanism* which is realized as a specific weight configuration in the synaptic weight space. From this perspective, *neural construction and adaptation processes* are the *heart* of any *scientific inductive process* in which a new theory is created or an already existing theory is adapted or changed.

Furthermore, it turns out that, as can be seen in Figure **2**, the “creation” or construction of new theories or knowledge does not bring forth “really new” knowledge. Rather, the context of discovery can be described as a *search process* in an already *predetermined space of possible theories*. This space is predetermined by the *paradigm* (Kuhn 1970) which has been chosen by the cognitive system. The *goal* of this search process is to *optimize* the *fit* and the level of consistency within the boundaries of this paradigm. Most research which is done in modern natural sciences turns out to be optimizing sets of parameters, methods, experimental set-ups, etc. leading to a better fit and consistency between predicted and actual phenomena. This may sound rather disappointing and provocative in the context of the epistemological and social status which scientific knowledge or theories normally claims to have. As an implication of the “cognitive view” as well as of a historic view of scientific processes, the notion of ultimate, objective, or true knowledge has to be seriously questioned. The constructivist perspective suggests to replace it by the concepts of *system relativity*, *functional fitness*, and *viability*.

1.6.2 Constructing new theory spaces: paradigmatic shifts and genetic space

In the picture about embedding scientific processes in neural dynamics, which has been presented so far, a couple of questions remain unanswered: What happens, if the search in a particular theory space is not successful or unsatisfactory? What happens, if the cognitive system’s goals and desires change? Who defines the (semantics of the) dimensions of the theory space? Which role does the *paradigm* play, and what happens, if the scientist (alias cognitive system) changes the paradigm or

---

[3] In many cases (e.g., particle accelerators, a laboratory situation in a psychological or biological experiment, etc.) one goes even so far as to *manipulate* the environmental constraints in order to increase the level of functional fit.

“invents”/constructs a new paradigm? As an implication from section 1.6.1 one can understand that what Kuhn (1970) refers to as a *paradigm* has its cognitive foundation in the structure of a particular synaptic weight space – note that its is not a certain configuration of synaptic weights, but the general architecture which opens up a space of possible weight configurations, and, thus, a space of possible theories which is searched through in the course of learning and adaptation (= “normal science”).

From that arises the interesting question, what a paradigmatic shift means in this cognitive perspective, and which mechanisms could be responsible for this phenomenon. As is well known from the history of science, in most cases a paradigmatic shift leads to completely new concepts, perspectives, and categories of how to view and understand a certain environmental phenomenon. How do these new categories, terminologies, and theories emerge on a cognitive level? In order to approach this problem from a cognitive perspective, it is necessary to have a closer look at the interactions between evolutionary processes and ontogenetic dynamics in the neural representational substratum.

Figure 3: The relationship between the genetic space, the synaptic weight space, the activation space, the behavioral dynamics/strategies, and the environmental structure.

Abstractly speaking we are confronted with *three degrees of freedom* in the representational dynamics of a cognitive system: (i) in the evolutionary dynamics a particular genetic code is “selected” and instantiated. In the process of genetic *expression* and interaction with the environment an organism develops. This organism has a (neural) representation system at its disposal. (ii) The second degree of freedom consists in the representational dynamics having been discussed above (i.e., learning, neural plasticity, search in weight space, etc.). The current state of the representation/weight space gives rise to a structure/dynamics in activation space. (iii) The dynamics of spreading activations instantiates states in the activation space and leads to the externalization of behavior. The success or failure of this behavior (= level of functional fitness of the behavior and of the representational structure) causes changes in the representational dynamics via learning/adaptation mechanisms. Over more generations the success or failure of the basic architecture of the representation system, the resulting behavior, as well as the genetically encoded basic body structures, developmental instructions, and learning/adaptation mechanisms cause a genetic drift (“evolution”). More abstractly speaking, the genetic code changes over time and gives rise to a newly structured representational space (= synaptic weight space, potential space of possible representational configurations of an organism) and, thus, to a potentially completely new set of behavioral strategies (see Figure **3**). In the course of ontogenesis the synaptic weight space is searched through. What can these interactions between evolutionary and ontogenetic representational dynamics teach us for our problem of paradigmatic shifts?

In section 1.6.1 it has been shown that – on a cognitive level – “normal science” (in the sense of Kuhn (1970)) can be characterized as a *search* and *optimization process* in theory space. It is embedded in the adaptation, learning, and construction processes of the whole neural representation system in the synaptic weight space. The goal is to find consistency between the environment and the theories which are generated (i.e., moving points in theory space) within the context and boundaries of the chosen paradigm. In this terminology a paradigmatic shift can be described as the construction of a completely new theory space. It consists of different dimensions, new and different semantics in the dimensions, and different representational and behavioral dynamics. In order to test this potential space of new theories, this new theory space has to be explored, as described in section 1.6.1 (“puzzle solving”). Looking at examples from history of science, one can see that the introduction of new paradigms often brought about some kind of *surprise* about the new way of looking at and structuring well known phenomena.

It is this “irrational” and unexpected character which makes paradigmatic shifts so interesting. In “normal science” (Kuhn 1970) most results are rather predictable and the theories being responsible for them have to undergo only minor adaptations in the course of a search process in a predetermined space of possible theories. Contrary to already established paradigms, newly constructed and unexplored paradigms are in most cases based on completely new concepts, basic assumptions, terminologies, and methods. This “irrational” character suggests that the (cognitive) processes being involved in *generating paradigmatic shifts* might have *evolutionary* character: a new paradigm is brought forth in a trial-&-error manner. It is even more hypothetical than the generation of a new theory in the context of an already established theory space/paradigm. This is due to the fact that at the moment of the conception of a new paradigm only a completely new hypothetical framework and space of potential theories is generated/suggested. Hence, there is rather high risk involved in this process. It can be compared to the process of expressing a gene which has undergone some kind of mutation. It is completely unclear, whether the resulting organism and its potential representational structures and behaviors will be capable of surviving. Similarly, at the moment of the conception of a new paradigm, a totally new potential theory space is created which has to be explored by the process of “normal science” – it is not at all clear, whether this space of potential theories will be successful or not. The mechanisms being involved in paradigmatic shifts have a lot in common with evolutionary operators which are applied to cognitive/representational structures: the introduction of completely new and unexpected categories, making use of metaphors, combining aspects from different theories, etc. have a lot in common with random mutations, cross over operators, etc. By applying these operators, a completely new theory space is established in which the key with the maximum functional fit with respect to a specific goal (e.g., of prediction or control) is searched (see also Figure **3**). From this perspective it is also clear that two or more different theories can account for the same phenomenon – two keys may account for a specific structure of the environment. As the goal is *not* to create an image or 1:1 mapping of the environment, but to construct correlations and consistencies in the form of functionally fitting behavior, it is no contradiction that two or more theories can account for the same phenomenon by making use of different representational categories. In any case, the interaction between evolutionary mechanisms and ontogenetic representational dynamics could shed some light on the mysterious phenomenon of paradigmatic shifts in science. Evolutionary operators act as “paradigm generators”; each of these paradigms establishes a space of potential theories which has to be searched through by applying the learning/adaptation strategies having been presented above.

Critics of such an evolutionary perspective of growth and development of scientific knowledge (e.g., Thagard 1988) are right in stating that a purely blind search for new scientific concepts is not an adequate model. That is why the focus has not only been on phylogenetic processes, but also on ontogenetic learning/adaptation including a feedback with the environmental dynamics. The important point is the interaction between the directed ontogenetic and neurally based learning, search, adaptation, and construction processes and the "blind" phylogenetic processes. Evolutionary variation "blindly" brings forth a completely hypothetical space of knowledge/strategies (paradigm) which is explored in a more or less directed manner in the course of ontogenetic development. In this process of exploration the new paradigm will prove its in-/adequacy very soon by receiving a feedback from the interaction with the environment. Contrary to Thagard's view(1988) that "the biological roots of the human information processing system are not directly relevant to the task of developing a model for the growth of scientific knowledge" (p 105), the presented concepts suggest that so-called scientific processes are not at all abstract processes occurring in a detached system called science. Rather, they are embedded in the dynamics of neural systems and are the result of heavy interaction between a group of neural systems and the environment – thus, it is no wonder that the process of science follows a structurally similar dynamics as cognitive processes.

Instead of a conclusion: Does a purely positivistic and cognitive science account cover the "whole story" about science?

Although both science and cognitive systems have some kind of representation of the world as their goal, it follows from the line of arguments that (scientific or common-sense) knowledge is reduced to playing the role of a *tool* or *instrument* aiming at mastering the world. Looking at the achievements of (natural) science during the last centuries, one has to admit that they have been extremely successful and have developed powerful theoretical as well as physical tools which are capable of (intellectually and physically) mastering/controlling large parts of the environmental dynamics. It seems that the process of science and its understanding having been developed throughout this paper has not only been successful, but also has gained a deep impact on our society, on our thinking, and especially on how we perceive and understand the world. What makes scientific theories so powerful and at the same time "seducing" is not only their capacity of explaining and predicting certain phenomena, but their *efficiency* concerning *manipulation* and *control* of the environmental dynamics. This has lead to an understanding and view of the world which is very much determined by the categories of efficiency, control, applicability, and functionality. It seems that these categories find their "epistemological justification" in the claims being made by the constructivist framework and by the concepts having been discussed in sections 0f. This fact is – per se – not a bad development, but it has a couple of serious implications on an epistemological level as well as in the realm of the (common-sense) view/understanding of the world:

If scientific theories (and knowledge in general) aim primarily at the manipulation and control of a certain reality, the original goal of achieving a *profound understanding* of the phenomenon under investigation has been more or less replaced by developing knowledge which can be (technologically) applied in one way or the other. As a result of our discussion above, it suffices to construct knowledge/theories about the (exterior) functional relationships of a particular environmental dynamics in order to achieve and satisfy this goal of prediction and control. The classical understanding of science and philosophy, however, did not content itself with reducing its understanding of a reality to describing its exterior functional relationships, but tried to go further in order to achieve a more profound understanding of a phenomenon "*from within*". From an epistemological perspective this means that studying and searching for a phenomenon per se has been replaced by a purely *functionalist* approach in modern science. This gives rise to the *seducing* assumption that, if one is able to successfully predict and control a particular part of reality, one possesses also a profound understanding of its meaning in all of its dimensions. Such an approach to reality is surely important and helpful, if we are interested in controlling it, but it remains an open question, whether it is intellectually satisfying for our mind…

Some of the reasons for this "silent dissatisfaction" have been touched already above. One of the main reasons seems to be the reduction of environmental phenomena (e.g., of life or cognition) to their *functional*, *relational*, *measurable* and *quantitative* aspects (which is not only correct, but also necessary out of the perspective and methodology of the natural sciences). The question is, however, whether we are covering the whole story, when reality is seen only under the aspect of being *measurable* or computable. The problem seems to be that natural sciences pretend and claim to know *what* a certain phenomenon, such as life or cognition, is by taking only the *measuring* or *computation* of *certain effects* of these phenomena as their point of departure for their investigation and analysis. In other words, the *effects* and their descriptions are *confused* with the investigated phenomenon *itself*.

This *quantitative analysis* of these phenomena remains at the material surface; such an approach is only capable to explain the *physical conditions* and the "*how*" of the object under investigation. This leads to a dissatisfaction having its roots in the fact that the object of any scientific investigation is not so much the (meaning of the) phenomenon itself, but its functional relationships being expressed in a rule-like form in a theory. Our intellect, however, is not only made for discovering the functioning/"how" of a given reality (be it a stone or a highly complex organism), but also – and this seems to be its finality – for investigating "*what*" a given reality is/means in its most profound sense. This question concerns the distinction between the *exercise* ("how" does a realty function/react/…?) and the *determination* ("what" is this reality?) of a reality which is often confused in the natural sciences: if one is studying only the dynamics or behavioral patterns of a phenomenon – this is what natural sciences are mainly doing –, one will never arrive at understanding its deeper nature and determination. This leads to this "silent dissatisfaction" having been mentioned above. The result are sophisticated (if-then-) relationships, but this kind of knowledge seems to stick at the exterior surface of the observed phenomenon. Form our experience we know that our thinking is capable of and longing for going further: namely, to get to know the reality "from within", from its interior, in its determination.

Furthermore, it seems to be logical that one can arrive only at an adequate understanding of a reality, if he/she knows (first) *what this reality is as it is* (and *not* as something as being measured or computed). Only, if one accepts this *primacy of determination*, one will be able to penetrate deeper and understand more adequately this phenomenon by studying its "how". This does not mean that the knowledge stemming from natural sciences is irrelevant for a more profound understanding, but one has to be aware that it covers only a rather *small fraction* of a given reality;

namely, the *reality as far as it is measurable or computable*. This becomes especially evident, if the given reality is a living being or, even more, if it is a human person and its cognitive abilities.

It seems to be inherent to the constructivist framework to be closer to the side of the “how”. This is due to having its roots in part in modern science, cybernetics, and systems theory which are mainly concerned with these aspects of reality. Hence, it is no wonder that the approach of embedding science in cognitive processes having been presented in sections 0f might explain the functioning of science quite well, but it does not really fully satisfy our intellect for the reasons having been brought forth above. A kind of circular explanatory strategy seems to have sneaked in. If the constructivist approach claims to provide a conceptual and epistemological foundation for a more adequate understanding of the world, and more specifically, for the process of science in all its (cognitive) dimensions, it will be necessary to shift the focus and take more seriously the issues having been addressed above as well.

References

Arbib, M.A. (Ed.) (1995). *The handbook of brain theory and neural networks*. Cambrdge, MA: MIT Press.

Ashby, R. W. (1964). *An introduction to cybernetics*. London: Methuen.

Bechtel, W. (1988). *Philosophy of science. An overview for cognitive science*. Hillsdale, N.J.: L.Erlbaum.

Belew, R.K. (1990). Evolution, learning, and culture: computational metaphors for adaptive algorithms. *Complex Systems 4*, 11—49.

Belew, R.K., J. McInerney, and N.N. Schraudolph (1992). Evolving networks: using the genetic algorithm with connectionist learning. In C.G. Langton, C. Taylor, J.D. Farmer, and S. Rasmussen (Eds.), *Artificial Life II*, Redwood City, CA, pp. 511—547. Addison-Wesley.

Berger, P. and M. Singer (1992). *Dealing with genes: the language of heredity*. Mill Valley, CA: University Science Books.

Brakel, J.v. (1994). Cognitive scientism of science. *Psycoloquy (electronic journal) 5*(20). (filename: scientific-cognition.3.vanbrakel).

Brown, T.H., A.H. Ganong, E.W. Kariss, and C.L. Keenan (1990). Hebbian synapses: biophysical mechanisms and algorithms. *Annual Review of Neuroscience 13*, 475—511.

Cangelosi, A., D. Parisi, and S. Nolfi (1994). Cell division and migration in a genotype for neural networks. *Network: computation in neural systems 5*(4), 497—516.

Churchland, P.M. (1979). *Scientific realism and the plasticity of mind*. Cambridge, New York: Cambridge University Press.

Churchland, P.M. (Ed.) (1989). *A neurocomputational perspective — the nature of mind and the structure of science*. Cambridge, MA: MIT Press.

Churchland, P.M. (1991). A deeper unity: some Feyerabendian themes in neurocomputational form. In G. Munevar (Ed.), *Beyond reason: essays on the philosophy of Paul Feyerabend*, pp. 1—23. Dordrecht, Boston: Kluwer Academic Publishers. (reprinted in R.N.Giere (ed.), Cognitive models of science, Minnesota Studies in the Philosophy of Science XV, 1992).

Churchland, P.M. (1995). *The engine of reason, the seat of the soul. A philosophical journey into the brain*. Cambridge, MA: MIT Press.

Churchland, P.M. and P.S. Churchland (1990). Could a machine think? *Scientific American 262*(1), 32—37.

Churchland, P.S. and T.J. Sejnowski (1992). *The computational brain*. Cambridge, MA: MIT Press.

Edelman, G.M. (1988). *Topobiology: an introduction to molecular embryology*. New York: Basic Books.

Essen, D.C.v., C.H. Anderson, and D.J. Felleman (1992). Information processing in the primate visual system: an integrated systems perspective. *Science 255*, 419—423.

Fodor, J.A. (1975). *The language of thought*. New York: Crowell.

Fodor, J.A. (1981). *Representations: philosophical essays on the foundations of cognitive science*. Cambridge, MA: MIT Press.

Foerster, H.v. (1973). On constructing a reality. In W.F. E. Preiser (Ed.), *Environmental design research*, Volume 2. Stroudsburg, PA: Hutchinson & Ross. (reprinted in P.Watzlawick (ed.), The invented reality, Norton, pp 41—61, 1984).

Gazzaniga, M.S. (Ed.) (1995). *The cognitive neurosciences*. Cambridge, MA: MIT Press.

Giere, R.N. (Ed.) (1992). *Cognitive models of science*, Volume XV of *Minnesota studies in the philosophy of science*. Minneapolis: University of Minnesota Press.

Giere, R.N. (1994). The cognitive structure of scientific theories. *Philosophy of Science 61*, 276—296.

Glasersfeld, E.v. (1984). An introduction to radical constructivism. In P. Watzlawick (Ed.), *The invented reality*, pp. 17—40. New York: Norton.

Glasersfeld, E.v. (1991). Knowing without metaphysics. Aspects of the radical constructivist position. In F. Steier (Ed.), *Research and reflexivity*, pp. 12—29. London; Newbury Park, CA: SAGE Publishers.

Glasersfeld, E.v. (1995). *Radical constructivism: a way of knowing and learning*. London: Falmer Press.

Hebb, D.O. (1949). *The organization of behavior; a neuropsychological theory*. New York: Wiley.

Heiden, U.a.d. (1992). Selbstorganisation in dynamischen Systemen. In W. Krohn and G. Küppers (Eds.), *Emergenz: die Entstehung von Ordnung, Organisation und Bedeutung*, pp. 57—88. Frankfurt/M.: Suhrkamp.

Hertz, J., A. Krogh, and R.G. Palmer (1991). *Introduction to the theory of neural computation*, Volume 1 of *Santa Fe Institute studies in the sciences of complexity. Lecture notes*. Redwood City, CA: Addison-Wesley.

Hinton, G.E. (1987). Connectionist learning procedures. Technical Report CMU-CS-87-115, Carnegie-Mellon University, Pittburgh, PA.

Holland, J.H. (1975). *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. Ann Arbor: University of Michigan Press.

Horgan, T. and J. Tienson (1996). *Connectionism and the philosophy of psychology*. Cambridge, MA: MIT Press.

Hubel, D.H. (1988). *Eye, brain, and vision*. New York: Freeman.

Hubel, D.H. and T.N. Wiesel (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology 195*, 215—243.

Jessel, T.M. (1991). Neuronal survival and synapse formation. In E.R. Kandel, J.H. Schwartz, and T.M. Jessel (Eds.), *Principles of neural science* (3rd ed.)., pp. 929—944. New York: Elsevier.

Kandel, E.R., J.H. Schwartz, and T.M. Jessel (Eds.) (1991). *Principles of neural science* (3rd ed.). New York: Elsevier.

Kosso, P. (1992). *Reading the book of nature. An introduction to the philosophy of science*. Cambridge: Cambridge University Press.

Kuhn, T.S. (1970). *The structure of scientific revolutions* (2nd ed.). Chicago: University of Chicago Press.

Langton, C.G. (Ed.) (1989). *Artificial Life*, Redwood City, CA. Addison-Wesley.

Langton, C.G. (Ed.) (1994). *Artificial Life III*, Redwood City, CA. Addison-Wesley.

Langton, C.G. (Ed.) (1995). *Artificial Life. An Introduction*. Cambridge, MA: MIT Press.

Maturana, H.R. (1991). Science and daily life: the ontology of scientific explanations. In F. Steier (Ed.), *Research and reflexivity*, pp. 30—52. London; Newbury Park, CA: SAGE Publishers.

Maturana, H.R. (Ed.) (1998). *Biologie der Realität*. Frankfurt/M.: Suhrkamp.

Maturana, H.R. and F.J. Varela (Eds.) (1980). *Autopoiesis and cognition: the realization of the living*, Volume 42 of *Boston studies in the philosophy of science*. Dordrecht; Boston: D.Reidel Pub. Co.

Miller, G., P. Todd, and S. Hedge (1989). Designing neural networks using genetic algorithms. In J.D. Schaffer (Ed.), *Proceedings of the Third International Conference on Genetic Algorithms*, San Mateo, CA. M. Kaufmann Pub.

Mitchell, M. and S. Forrest (1994). Genetic algorithms and artificial life. *Artficial Life 1*(3), 267—291.

Newell, A. (1980). Physical symbol systems. *Cognitive Science 4*, 135—183.

Newell, A. and H.A. Simon (1976). Computer science as empirical inquiry: symbols and search. *Communications of the Assoc. for Computing Machinery (ACM) 19*(3), 113—126. (reprinted in M.Boden (ed.), The Philosophy of Artificial Intelligence, Oxford University Press, 1990; in German in D.Münch (ed.), Kognitionswissenschaft, Suhrkamp, 1992).

Peschl, M.F. (1994a). Embodiment of knowledge in the sensory system and its contribution to sensorimotor integration. The role of sensors in representational and epistemological issues. In P. Gaussier and J.D. Nicoud (Eds.), *From perception to action conference*, Los Alamitos, CA, pp. 444—447. IEEE Society Press.

Peschl, M.F. (1994b). *Repräsentation und Konstruktion. Kognitions- und neuroinformtische Konzepte als Grundlage einer naturalisierten Epistemologie und Wissenschaftstheorie*. Braunschweig/Wiesbaden: Vieweg.

Peschl, M.F. (1997). The Representational Relation Between Environmental Structures and Neural Systems: Autonomy and Environmental Dependency in Neural Knowledge Representation. *Nonlinear Dynamics, Psychology, and Life Sciences 1*(2), 99—121.

Port, R. and T.v. Gelder (Eds.) (1995). *Mind as motion: explorations in the dynamics of cognition*. Cambridge, MA: MIT Press.

Roth, G. (1992). Kognition: die Entstehung von Bedeutung im Gehirn. In W. Krohn and G. Küppers (Eds.), *Emergenz: die Entstehung von Ordnung, Organisation und Bedeutung*, pp. 104—132. Frankfurt/M.: Suhrkamp.

Roth, G. (1994). *Das Gehirn und seine Wirklichkeit. Kognitive Neurobiologie und ihre philosophischen Konsequenzen*. Frankfurt/M.: Suhrkamp.

Rumelhart, D.E. and J.L. McClelland (Eds.) (1986). *Parallel Distributed Processing: explorations in the microstructure of cognition. Foundations*, Volume I. Cambridge, MA: MIT Press.

Schmidt, S.J. (1998). *Die Zähmung des Blicks. Konstruktivismus — Empirie — Wissenschaft*. Frankfurt/M.: Suhrkamp.

Schwartz, E.L. (Ed.) (1990). *Computational neuroscience*. Cambridge, MA: MIT Press.

Sejnowski, T.J., C. Koch, and P.S. Churchland (1988). Computational neuroscience. *Science 241*(4871), 1299—1306.

Steels, L. (1995). The Artificial Life Roots of Artificial Intelligence. In C.G. Langton (Ed.), *Artificial Life. An Introduction*, pp. 75—110. Cambridge, MA: MIT Press.

Steier, F. (Ed.) (1991). *Research and reflexivity*. London; Newbury Park, CA: SAGE Publishers.

Thagard, P. (1988). *Computational philosophy of science*. Cambridge, MA: MIT Press.

Varela, F.J., E. Thompson, and E. Rosch (1991). *The embodied mind: cognitive science and human experience*. Cambridge, MA: MIT Press.

Wiener, N. (1948). *Cybernetic. Control and communication in the animal and the machine*. New York: Wiley.

Winston, P.H. (1992). *Artificial Intelligence* (3rd ed.). Reading, MA: Addison-Wesley.

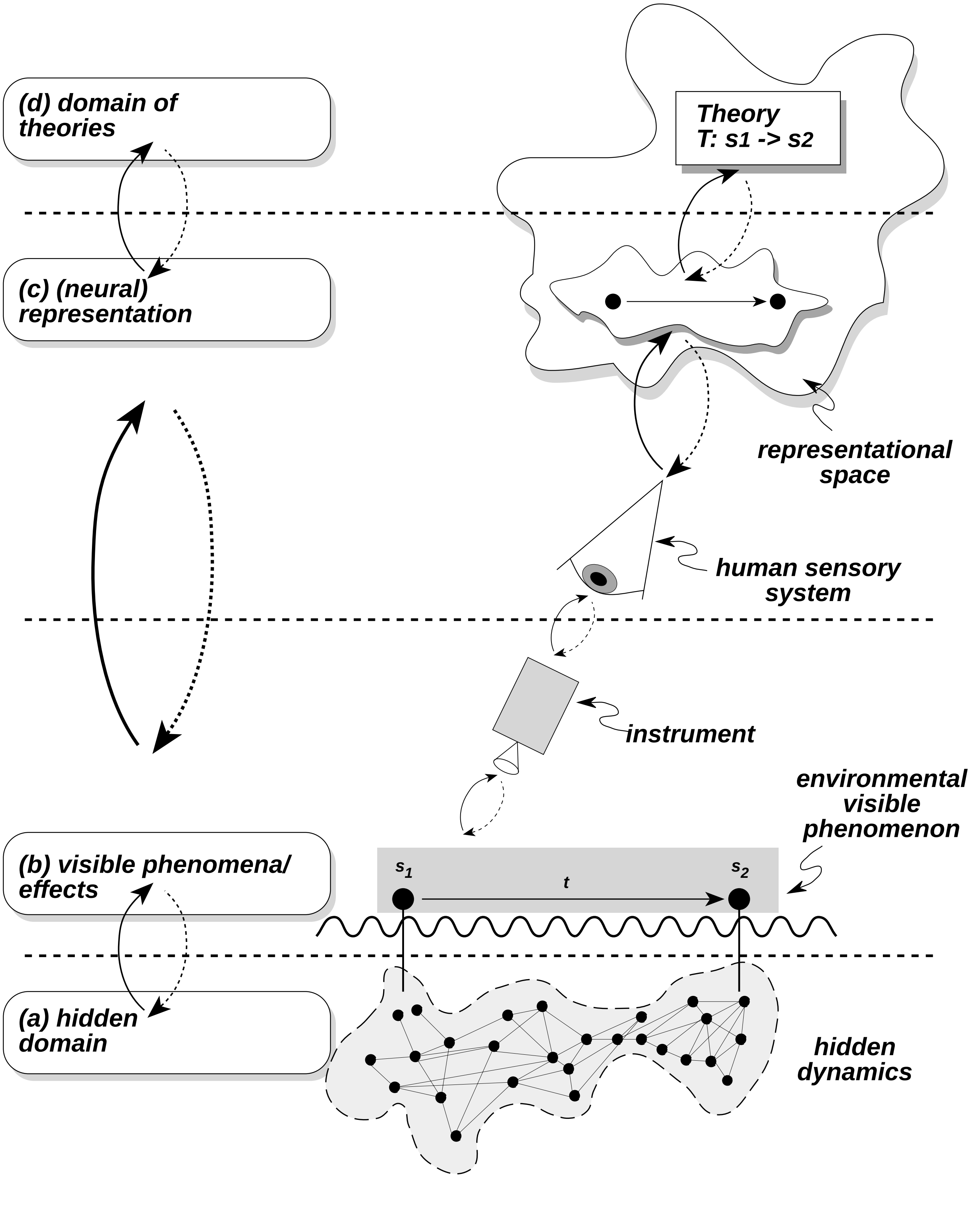

(d) domain of theories
(c) (neural) representation
(b) visible phenomena/ effects
(a) hidden domain
Theory
T: s1 -> s2
representational space
human sensory system
instrument
environmental visible phenomenon
s1
t
s2
hidden dynamics

Figure 1

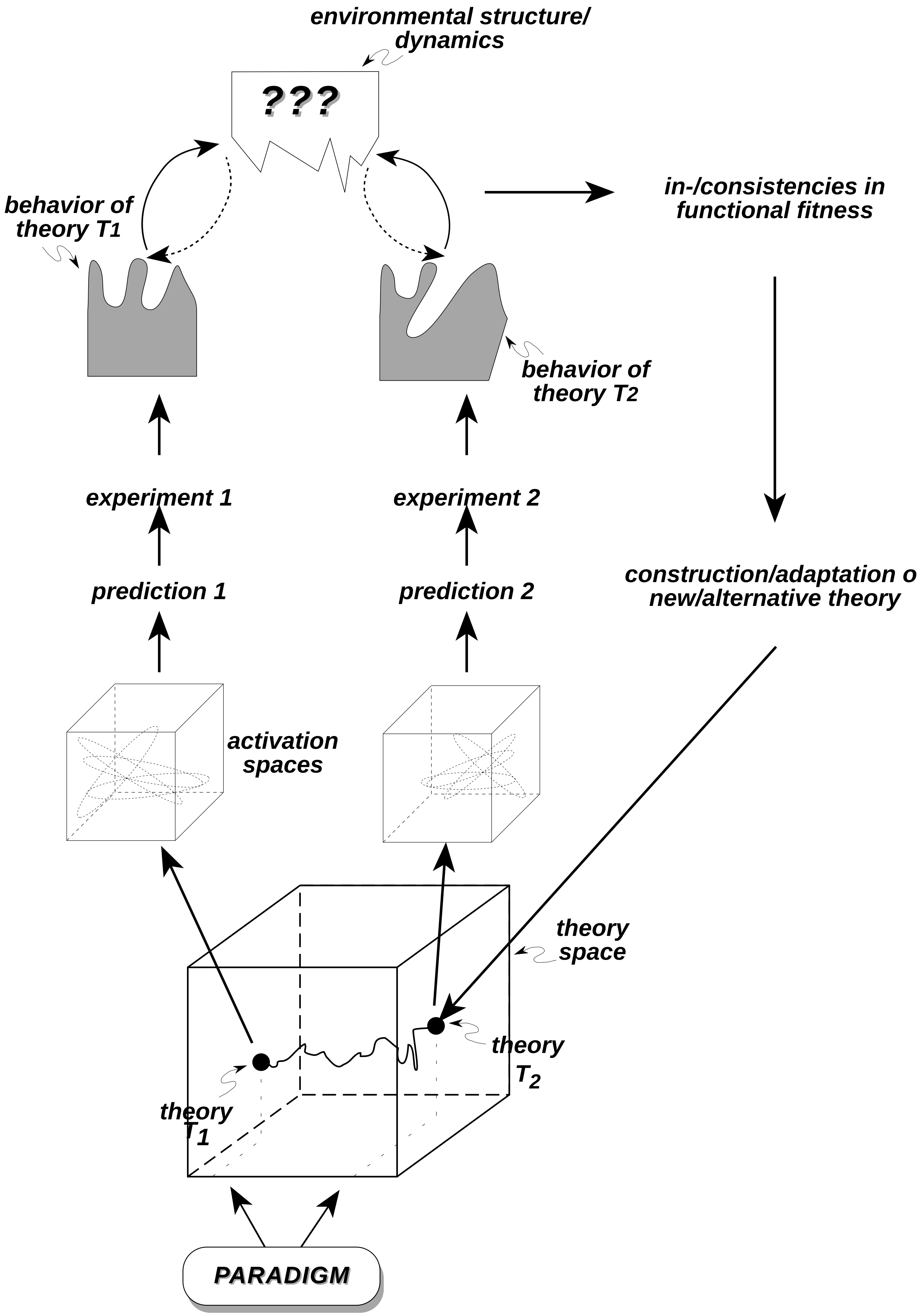
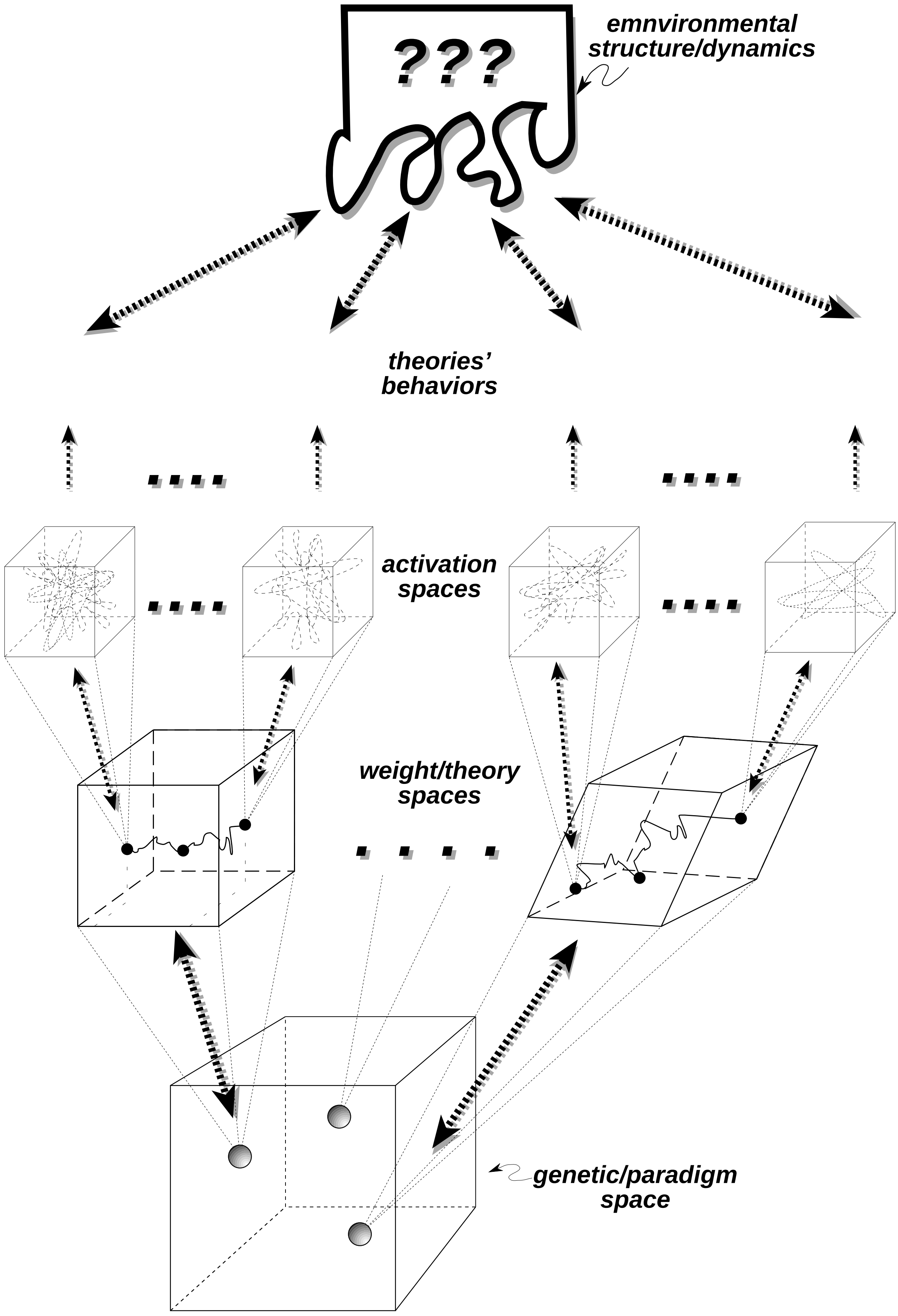

environmental structure/
dynamics
???
behavior of
theory T1
in-/consistencies in
functional fitness
behavior of
theory T2
experiment 1
experiment 2
construction/adaptation o
new/alternative theory
prediction 1
prediction 2
activation
spaces
theory
space
theory
T2
theory
T1
PARADIGM

Figure 2



emnvironmental
structure/dynamics
???
theories'
behaviors
activation
spaces
weight/theory
spaces
genetic/paradigm
space

Figure 3