Part XI

IMPLICATIONS OF THE COGNITIVE SCIENCES FOR PHILOSOPHY OF SCIENCE

https://doi.org/10.1086/psaprocbienmeetp.1990.2.193085 Published online by Cambridge University Press

Implications of the Cognitive Sciences for the Philosophy of Science[1]

Ronald N. Giere

University of Minnesota

1. Introduction.

Does recent work in the cognitive sciences have any implications for theories or methods employed within the philosophy of science itself? The answer to this question depends first on one's conception of the philosophy of science and then on the nature of work being done in the various different fields comprising the cognitive sciences. For example, one might think of the philosophy of science as being an autonomous discipline that is both logically and epistemologically prior to any empirical inquiry. If the cognitive sciences are empirical sciences, then research in the cognitive sciences could not have any significant implications for the philosophy of science. And that would be the end of the matter. Logical Empiricism is now typically understood as having exemplified this point of view.

More specifically, Logical Empiricism took it for granted (i) that scientific knowledge should be understood as ideally having the structure of a formal logical calculus, and (ii) that the empirical warrant for scientific claims is given by directly observed data together with formal rules which determine the weight of the evidence for or against the particular claims in question. In a Logical Empiricist framework, therefore, scientific knowledge is structured linguistically and the epistemological relationship between evidence and theory is a linguistic relationship. Actual human cognition, as investigated by the cognitive sciences, is irrelevant to such relationships. This was part of the legacy of Frege, who preached forcefully against the sin of psychologism. One of the main messages of this paper is that it is finally time to put this legacy in its proper perspective, which is, that it is not particularly useful for understanding the nature of science.

The situation is not much different if one adopts the picture of the philosophy of science developed by philosophical critics of Logical Empiricism such as Lakatos (1970) or Laudan (1977, 1984). Both agree with Logical Empiricists that theories are linguistic structures. They disagree only that there is any quasi-logical epistemological relationship between data and individual theories. For Lakatos, Laudan, and their followers, appraisal applies only to a series of theories. The question for both is whether the whole series of theories is "progressive" or not, where progress is measured either in

PSA 1990, Volume 2, pp. 419-430

terms of predicted novel empirical content or problem solving effectiveness. The traditional connections between rationality and truth (or probability) are broken. Rationality is reduced to a kind of progress, but the proposed kinds of progress have little to do with actual human cognition. So here again there can be no substantial implications for the philosophy of science from the direction of the cognitive sciences.

Things appear quite different from the viewpoint developed by Kuhn thirty years ago in *The Structure of Scientific Revolutions* (1962) and in subsequent philosophical writings (1977). For Kuhn, scientific knowledge is not adequately captured by laws, theories, or other linguistic structures. Nor is there any epistemic relationship between linguistically formulated theories and descriptions of evidence. Rather, scientific knowledge is embodied in a scientific community which shares a family of "exemplars" — specific examples which are judged to exhibit solutions to important problems. There is no epistemic relationship which abstracts from the actual judgments of a scientific community about the exemplary status of a family of specific examples, or about the applicability of certain of these exemplars to a new problem.

Here there is ample room for input from the cognitive sciences. Relevant inquiry into the nature of exemplars and of judgments about exemplars is the kind of thing one could, at least in principle, expect to find being studied by cognitive scientists. Kuhn himself clearly recognized this possibility with his original appeal to gestalt psychology and subsequent appeals to cognitive psychology and psycholinguistics (Kuhn 1977).

With these examples in mind, one can make a stab at formulating some conditions on any conception of the philosophy of science which could accommodate the relevance of research in the cognitive sciences. (i) The philosophy of science must be *naturalistic* in the sense that its claims are subject to test by empirical data in the same way that ordinary scientific claims are so subject. This is hardly sufficient, however, since Laudan (1977) claims his theory of science to be naturalistic in that it is to be authenticated by appeal to the history of science. So we must add either (ii) understanding the nature of scientific knowledge, its laws, theories, or whatever, requires understanding the cognitive capacities and activities of scientists; or (iii) judging the epistemic merit of scientific knowledge claims requires appeal to the judgmental capacities and activities of scientists; or both (ii) and (iii).

Now what about the cognitive sciences? What implications have they for a "cognitive" philosophy of science satisfying the above conditions? Within the cognitive sciences, it is common to distinguish three overlapping disciplinary clusters which tend to be thought of as providing three different *levels* of analysis: (i) cognitive neuroscience, (ii) cognitive psychology, and (iii) artificial intelligence (AI). The standard view is that the functional units get larger and more abstract as one moves up the hierarchy from neuroscience to AI. The implications for the philosophy of science are different, and sometimes conflicting, depending on which cluster of disciplines one examines. I shall consider all three in order of increasing abstractness.

2. Implications from Neuroscience.

The person who has done most to exploit recent developments in neuroscience in the service of the philosophy of science is Paul Churchland (1989). For Churchland, the brain is a network consisting of layers of neuron-like units. At each moment, every unit exhibits an activation level which it transmits to units in the next "higher" layer. More importantly, each pathway from one unit to units in the next higher layer is characterized by a "weight;" which may be positive or negative, that regulates the

https://doi.org/10.1086/psaprocbienmeetp.1990.2.193085 Published online by Cambridge University Press

relative strength of the signal transmitted along that particular pathway. Thus, the strength of the signal coming into any higher level unit is a weighted sum of the activation levels of the units which feed into that unit. If we assume that the activation levels of units in the initial layer are determined by sensory inputs, then the activation levels of units at all higher layers are thereafter determined by the set of weights which regulate the strength of signals transmitted from unit to unit through the various levels. The output would be the set of activation levels at some higher layer of units.

An exemplar of this kind of network is a simple three-layered network designed to distinguish rocks from mines in the bottom of a harbor (1989, pp. 164-69). The input layer consists of 13 units each representing the average energy in a range of the frequency spectrum of a sonar signal. Each of these units is connected, with a given relative weight, to each of seven units in the second, so-called "hidden layer." Similarly, each of these seven units is connected, again with given weights, to both of two output units, representing "rock" and "mine " respectively. A resultant activation level of near one for the rock unit and near zero for the mine unit means that the network has processed the input as being a signal from a rock rather than a mine.

This network is "trained up" by being fed examples of sonar spectra from known objects. At first the network gives indeterminate answers. But after each example the weights are adjusted slightly so as to improve its answer. After a few thousand such training exercises, the network not only gives unambiguously correct responses for most of the original examples, but also for new examples. It has succeeded in isolating general features of mine and rock signals that permit reliable discrimination between the two.

If we say that the network has developed a *representation* of mine and rock signals, then this representation has some interesting features. First, the representation is not fundamentally propositional in nature. The network does not function by performing anything like logical operations on statements. There are no encoded statements on which to operate. Second, the representation is not localized anywhere in the network, but *distributed* throughout the network as a whole. Churchland argues that the representation is best characterized in terms of the total set of *weights* that regulate the propagation of activation strengths from the layer recording the sensory inputs to the output layer.

Neural networks function primarily as pattern recognizers. Thus, as Churchland points out, the human brain is particularly well-suited to implement a theory of science which takes Kuhnian exemplars as primary. The normal science activity of recognizing a new problem as being similar to an older, exemplary problem is just the kind of thing for which, on Churchland's view, scientists' brains ought to be particularly well adapted. This is an exciting, and, I think, fundamentally correct, insight.

Churchland goes on to suggest that we *identify* a scientist's theory with the weight vector that characterizes that scientist's neural network. This could apply to a scientist's whole neural network, a truly global theory, or just to some part, which would be a more local theory. In either case, it follows immediately that perception is necessarily theory laden. For perception just is the process of propagation up the network that yields a perceptual judgment, and that process is characterized by a particular weight vector.

Here I think that Churchland's reductionist proclivities have gotten the best of him. His enthusiasm for connectionist models has temporarily blinded him to the obvious fact that any adequate cognitive theory of science will require consideration of

https://doi.org/10.1086/psaprocbienmeetp.1990.2.193085 Published online by Cambridge University Press

cognitive structures at higher, and more abstract, levels of functional organization. This general point can be brought home with a few commonplace examples.

There is a whole subculture built around the activities of creating, producing, performing, marketing, and listening to music. Neuroscience no doubt has something to contribute to a theoretical understanding of the culture of music. But focusing solely on the synaptic weights of the brains of composers, performers, and listeners seems a hopeless way to pursue this subject. We need to be able to talk about individual musical works, types of music, styles of performance, and so on, at a higher level of organization.

Similarly, suppose that physicalism is true. That is, the whole world and everything in it, including all living things, is nothing more than one big quantum mechanical system. Nevertheless, attempting to pursue evolutionary biology at the quantum mechanical level would be a hopeless scientific enterprise. We cannot so easily dispense with talk of things like genotypes and founder populations. So also with an understanding of the culture and evolution of science as a cognitive activity.

One source of our disagreement may be that Churchland wants scientific theories to be solely in the brains of scientists. But this supposition ignores the obvious fact that scientists use a wide variety of "external" representational devices such as diagrams, graphs, and, of course, written words and equations. Maybe we should invoke an even more radical notion of distributed representations than that provided by neural networks. Even Kuhn's exemplars may have to be thought of not as being localized in the brains of individual scientists, but as distributed both among the brains of many scientists, and also among their many external representational devices.

It is understandable that Churchland should wish to solve all the major problems of the philosophy of science, like the theory ladenness of observation, at the level of neuroscience. Current connectionist models of the brain are impressive. But there is also a more modest role for neuroscience. Churchland himself gives a clear statement of this more modest role for epistemology in general:

> Making acceptable contact with neurophysiological theory is a long-term constraint on any epistemology: a scheme of representation and computation that cannot be *implemented* in the machinery of the human brain cannot be an adequate account of human cognitive activities. (1989, p. 156, italics added)

Implementation does not require the *identification* of higher-level functional entities with features of neuronal entities. It only requires that whatever is done with the higher-level entities be something than can actually be accomplished by flesh and blood scientists with flesh and blood brains. But Churchland's main point is unassailable. Being compatible with established results in neuroscience is a necessary condition for any adequate cognitive theory of science as a human activity.

3. Implications from Cognitive Psychology.

Among several good examples of research in the cognitive sciences being used to approach problems in the philosophy of science is Nancy Nersessian's study of the development of electrodynamics from Faraday to Einstein (1984, 1988, 1991). Unlike Lakatos (1970) or Laudan (1977), Nersessian is not concerned to show that conceptual change is rational in some special sense. And contrary to both classical Logical Empiricism and the original Kuhn, she argues that conceptual development in science is continuous but not cumulative. To develop an adequate theory of how such development might occur, she invokes ideas originating within cognitive psychology. These include theories of "mental models," like that of Johnson-Laird (1983), and of both analogical and imagistic thinking. Here I would like to highlight several general features of her account — features that I think should be part of any cognitive theory of science.

Nersessian argues that one cannot construct an adequate theory of conceptual change by focusing only on abstract structures — be they "concepts," "mental models," or whatever. The reason, I think, is that such structures possess no internal dynamical forces which could make changes of one type rather than another. They may contain structural possibilities for change, but they contain nothing to make changes happen.

The alternative, as Nersessian insists, is to think of conceptual change as something that is accomplished by cognitive agents. So a theory of conceptual change is necessarily a theory of the capacities and activities of cognitive agents. In my own terms (Giere 1989a), the basic units of analysis for a cognitive theory of science should be individual scientists — not concepts, theories, etc. So the problem of conceptual change becomes the problem of how scientists develop new conceptual structures.

Taking the individual scientist as the basic unit of analysis provides a natural bridge between the philosophy of science and both the history of science and the sociology of science. In spite of much recent historical research focusing on the social and institutional aspects of science, the individual scientist remains at the center of historical narratives. Similarly, recent sociology of science emphasizes the role of various types of human interests and human interactions in the development of particular sciences. But interests and human interactions can only influence the development of science in so far as they influence the thoughts and actions of individual scientists. Putting the scientist at the center of a cognitive theory of science thus makes it possible causally to connect a variety of interests and interactions with the actual historical course of science.

There are several possible ways of understanding projects like Nersessian's. One way is as an attempt to show that beginning with (i) a particular model (theory, or whatever) and (ii) the natural cognitive abilities of scientists, one can explain the development of a successor model. For example, to explain how Maxwell developed his field theory, one would need only his starting point with Faraday's theory plus an understanding of the relevant cognitive mechanisms at Maxwell's disposal.

This clearly is not Nersessian's project. I mention it only for comparison. In any actual scientific development, the later model would be dramatically underdetermined by the earlier model plus cognitive mechanisms. One would have to add at least some other models used by the scientist in question, as Maxwell used various mathematical models unknown to Faraday.

A second possible project requires creating a new, cognitive version of the internal-external distinction by claiming that only other "scientific" models need be considered, and not, for example, religious or political models. My own view would be that one cannot possibly legislate a priori what sorts of models one might have to consider in order to explain how a particular scientist got from the initial to the final set of scientific models. Any kind of model might play an important role. Robert Richards' (1987) study which suggests the influence of religious models on Darwin's thinking provides a recent example of this third sort of project.

A still more inclusive project would be one that includes not only other models, but motivations and interests as well. The recent sociological literature contains many cases for which it is claimed that one cannot explain how the new models came to be developed or accepted without invoking a variety of interests (Shapin 1982). If one understands the realm of the cognitive as excluding motivation and interest, then this fourth possible project would deny that there are always "cognitive" explanations of the development of new models. I would prefer to classify motivational factors as cognitive, so that then all explanations of new models might be appropriately cognitive. In any case, I suspect that this fourth project is more inclusive than many, including Nersessian, would prefer, but I doubt that anything less can do justice to the historical facts.

Given the origin of their concerns in the two decades following Kuhn's analysis of science, Nersessian and others have quite naturally focused on conceptual *change*. What is emerging from this work, I think, is an increasing realization that we lack a clear, widely shared theory of the nature of conceptual structures themselves. We are in the position of trying to develop a theory of conceptual change without having a good theory of concepts — the things that are supposed to change.

That we should be in this situation is not all that surprising. In their reaction to both Logical Empiricism and Kuhn, Lakatos and Laudan maintained the Logical Empiricist view that theories are fundamentally sets of statements. They sought to replace the *epistemological* doctrine of Logical Empiricism that there can be empirical support for individual theories with the doctrine that there can only be progress in a historical series of theories.

Most other attempts to talk about conceptual change likewise assumed a Logical Empiricist account of the nature of theories. The problem was to understand how one got from one theory, so characterized, to another. What Nersessian and others are coming to realize is that one cannot construct an adequate account of the process of theoretical change using a Logical Empiricist account of the nature of theories. One needs a better account.

It is not just that the Logical Empiricist account presupposes that it is adequate to think of a scientific theory as an axiomatic system formulated in first order logic. Rather it is that the Logical Empiricist account takes the basic representational relationship to be that of the *truth* of an individual *statement*. Nor is the later, more holistic, Quineian conception of theories much better. It merely moves us from individual statements to sets of statements. Truth remains the basic goal for representational success.

A "semantic", or "model theoretic", account of theories (Giere 1988; Suppe 1989; van Fraassen 1980, 1989) takes us a step in the right direction, but only, I now think, a small step. On this view, a model has the logical function of a predicate. Thus, rather than talking about statements being true or false, we talk about predicates being *true of* the world (or not). This is still too narrowly linguistic. We need a representational notion that is broad enough to encompass graphical, imagistic, and pictorial representations as well.

Here we find, I think, what is right now the most promising problem area in the philosophy of science for which the cognitive sciences might provide interesting solutions. Nor is it the case that the cognitive sciences lack candidates for the replacing of linguistic, or linguistic-like, structures as the basic representational devices of the sciences. Rather, there are too many candidate devices. It seems that every research area in the cognitive sciences employs a different notion. Thus the "mental models" of cognitive psychologists like Neisser (1976) and Johnson-Laird (1983) are different.

And these are different from the "mental models" of more developmentally oriented psychologists like Carey (1985) or Gentner (Gentner and Stevens 1983). And these are different again from the "mental models" of cognitive linguists like Lakoff (1987). The job for a cognitive philosopher of science, therefore, is not simply to find an adequate notion of representation to import from cognitive psychology. Rather, the task is to forge from these diverse conceptions a conception of representation in science that is adequate for constructing a good theory of scientific development. But the first step is finally to abandon the legacy of Frege and Russell and to realize that statements, or sets of statements, are just one type of representational device - and maybe even not the most important type employed in the actual practice of science.

Abandoning this tradition is a lot easier said than done. In addition to new theories of representation, we may need what used to be called a new metaphysics. In an age of naturalism this would be called a new "theoretical perspective." The old perspective was obtained by reading the structure of language into the world. So the world was thought of as consisting of states of affairs which mirror the structure of statements. A better perspective may be to begin with the world as something with many levels of complexity, so that it might be pictured from many angles. Different pictures may capture different aspects of the complexity, or different levels of the complexity. Simple questions of truth or falsity become irrelevant. But that does not mean that one is not genuinely representing the world.

I am inclined to take the metaphor of pictures quite seriously. Rather than taking representation by statements as fundamental, we should take the way in which pictures represent the world as fundamental. So there may be something to a picture theory of meaning after all, except that it is not statements themselves that picture the world. Rather, statements are just one type of device that may be used in constructing a picture, or model, of the world. It is the model that pictures the world. The problem, then, is to understand that relationship.

4. Implications from Artificial Intelligence.

More than most other disciplines within the cognitive sciences, artificial intelligence has been split by connectionism. It is here that the idea of a "second cognitive revolution" is most applicable. As is already clear from my earlier discussion of neuroscience, connectionism provides a way of linking the top and the bottom of the cognitive sciences hierarchy. Here I will confine my remarks to good old-fashioned, rule-based AI.

There are two sorts of traditional AI activities which have potential implications for a properly naturalized philosophy of science. One is the development of programs that can perform a variety of scientific tasks which go well beyond number crunching. These include: (i) Discovery programs, inspired by Herbert Simon and implemented by Pat Langley and others, which, using fairly general heuristics, can uncover significant regularities in various types of data. Programs such as BACON (Langley, et al 1987) and KAKEDA (Kulkarni and Simon 1988) provide exemplars of such programs. (ii) Programs which generate and evaluate causal models in the social sciences. Here the prototypes are the TETRAD programs developed by Clark Glymour and associates (1987). (iii) Programs which aid in the classification and resolution of anomalies arising in the course of theorizing and experimentation. Lindley Darden (1991) is currently developing such programs with particular reference to genetics.

Programs of this nature are potentially of great scientific utility. That potential is already clear enough to keep good people working on developing them further

(Shrager and Langley 1990). The implications of these sorts of programs for a cognitive philosophy of science are mainly *indirect.* The fact that they perform as well as they do can tell us something about the structure of the domains in which they are applied and about possible strategies for theorizing in those domains. In general, the role of these programs in science is more something to be explained by a cognitive philosophy of science than a resource to be deployed in developing such explanations.

Other philosophers of science, by contrast, have advocated philosophically much more ambitious projects for applying standard AI techniques to the philosophy of science. Paul Thagard (1988, 1989, 1991) provides a prominent example. Thagard explicitly seeks to distance himself from Logical Empiricism. (i) He regards his philosophy of science as a species of naturalized philosophy of science . (ii) He has a much more liberal view of the nature of scientific theories. For Thagard, theories are conceptual or propositional networks rather than axiomatic systems. (iii) Going beyond standard deductive and inductive logics, his model of scientific validation is based on a notion of explanatory coherence which includes such things as analogy.

Nevertheless, it still seems to me that the differences between Thagard's project and that of Logical Empiricism are at the level of *implementation* rather than of fundamental principles. For example: (i) Thagard maintains the fundamental view that scientific knowledge has a structure that can be adequately captured by some sort of propositional system. (ii) He maintains that the reasoning required in choosing one theory over another can be analyzed solely in terms of relationships among propositions. (iii) He advances a project for getting from the descriptive claims of a naturalistic philosophy of science to normative rules, such as inference to the best explanation. And he clearly intends that the normative force behind such rules be more than a matter of empirically based, means-end reasoning (1988, ch. 7).

Thagard would also like to think that his propositional structures are embodied in actual human thinking, thereby bridging the divide between AI and cognitive psychology. But so far there is little evidence that this is even possible, let alone actual. The interesting question is what Thagard would do in the face of empirical evidence that actual humans either can not or do not utilize anything like explanatory coherence in deciding among rival theories. But clear evidence one way or the other is likely to be a long time coming. We are left at the moment with a propositional structure whose normative claim on our allegiance remains obscure, and the hope that the minds of scientists embody such structures.

Thagard's writings contain two different approaches to the fundamental question of the nature of representation in science. His earlier account (1988) portrays scientific theories as a type of *production system.* In light of the well-known logical equivalence between production systems and axiomatic systems, Thagard argues that the equivalence pertains only to "expressive" power and not to "procedural" operations. That is, production systems are computationally more tractable and perhaps even psychologically more realizable than axiomatic systems. This useful distinction, under the labels "informational equivalence" and "computational equivalence" has long been employed by Herbert Simon (1978) to distinguish, for example, the difference between linguistic and pictorial representations (Larkin and Simon 1987).

In his more recent writings (1989, 1991), Thagard distinguishes conceptual systems from theories as sets of propositions utilizing the concepts from the corresponding conceptual system. He portrays conceptual systems as networks of localized concepts linked by part and kind relationships. For example, in the Ptolemaic conceptual system, the Sun and Mars are both planets, which are a kind of star, that is, wandering stars. In the Copernican conceptual system, by contrast, the Sun is a star, while the Earth and Mars are planets, which are now regarded not as stars but as satellites of the sun. Both systems deal with the same set of objects, but they are categorized differently.

The conceptual networks exhibited by Thagard and others seem to me to provide only a relatively superficial description of the endpoints of a scientific revolution. At most they provide ways of cataloguing scientific revolutions according to how conceptual networks get restructured. These networks provide no account whatsoever of the *dynamics* of conceptual development. Unlike Nersessian, Thagard does not even attempt to develop an account of how scientists, beginning with one conceptual system, construct a later conceptual system. For this project it would be better to begin with the *pictures* constructed by members of the rival camps. For example, a Ptolemaic picture would show the Earth at the center of a set of concentric circles, one of which represents the orbit of the Sun. A Copernican picture would have the Sun at the center with the orbit of the Earth represented by one of the concentric circles. The structure of Thagard's conceptual networks is embodied in these pictures, but the pictures included much more. And we know such pictures played a role in the actual thinking of participants at the time.

On Thagard's view (1989, 1991) the new conceptual system replaces the old because the new theory exhibits greater explanatory coherence relative to data shared by both theories. The overall explanatory coherence of any theory is a function of binary coherence relationships among statements of the theory together with statements of evidence. Thagard has developed a program, ECHO, which performs the requisite calculations of relative explanatory coherence for rival theories. The theoretical importance of his account, however, lies in the theory of explanatory coherence, which I cannot reproduce here. For those who wish to pursue the matter further, I offer the following caveats.

In the examples Thagard presents, such as the Copernican Revolution (Thagard 1991; Nowak and Thagard 1991), the respective theories are reconstructed by Thagard and his associates from classical texts. It is difficult to assess whether the process of reconstruction itself has introduced biases in favor of the known historical outcome. Even if it has not, the texts used were typically produced in a historical context in which the objectives were as much rhetorical as scientific. So Thagard may simply be analyzing the rhetorical structure of a text, not the reasoning of any of the participants (Giere 1989b). In any case, Thagard provides no independent evidence either that the participants made the choices they did because they perceived the greater explanatory coherence of the Copernican theory, or that their minds just naturally worked according to his principles.

A determination of relative explanatory coherence begins with two existing rival conceptual systems together with their corresponding theories. This account, therefore, assumes a version of the Logical Empiricist distinction between discovery and justification. But even assuming such a distinction, an account of the choice of one theory over a rival theory would have to include much more of the actual historical context, including the motivations and interests of the participants, than is embodied in the sorts of relationships among propositions countenanced in Thagard's theory of explanatory coherence. And it would have to include an understanding of experimentation as more than merely a means of producing evidence statements.

The differences between Thagard's approach and my own may be summarized by considering the difference in the labels we employ: *computational* philosophy of science versus *cognitive* philosophy of science. As I see it, a computational theory of science is one species of cognitive theory of science, one drawing primarily on the resources of traditional, rule-based AI. Those resources may be useful, but they seem to me far too limited for the task.

5. Conclusion.

In the wake of Kuhn's work, many philosophers of science managed to hold on to the basic framework of Logical Empiricism by marginalizing and assimilating Kuhn's insights as being concerned solely with the temporal development of scientific theories, a topic which simply was not on the agenda of Logical Empiricism. One finds now a similar reaction to the idea of importing theories and methods from the cognitive sciences into the philosophy of science. There is a tendency to marginalize and assimilate such studies as being concerned only with conceptual change, scientific discovery, or creativity.

My view is that, as was the case regarding Kuhn's historical critique, this is a defensive attempt to avoid facing the challenges to the fundamental assumptions of Logical Empiricism, particularly the assumption that the basic representational device in science is a statement, a set of statements, or some similar linguistic structure. The message coming from at least some of the cognitive sciences is that this simply is not so. The implication of these cognitive sciences for the philosophy of science is that the philosophy of science needs to be rethought from the ground up.

Note

[1]The author gratefully acknowledges the support of the National Science Foundation and the hospitality of the Wissenschaftskolleg zu Berlin.

References

Carey, S. (1985), *Conceptual Change in Childhood.* Cambridge: MIT Press.

Churchland, P.M. (1989), *A Neurocomputational Perspective*. Cambridge: MIT Press.

Darden, L. (1991), "Strategies for Anomaly Resolution". In *Cognitive Models of Science*, ed. R.N. Giere, Minnesota Studies in the Philosophy of Science, vol. 15. Minneapolis: Univ. of Minnesota Press.

Gentner, D. and Stevens, A.L. (1983), *Mental Models*. Hillsdale, NJ: Erlbaum.

Giere, R.N. (1988), *Explaining Science: A Cognitive Approach.* Chicago: University of Chicago Press.

______. (1989a), "The Units of Analysis in Science Studies". In *The Cognitive Turn: Sociological and Psychological Perspectives on Science*, ed. S. Fuller, M. DeMey, T. Shinn, and S. Woolgar, Sociology of the Sciences Yearbook. Dordrecht: D. Reidel.

______. (1989b), "What Does Explanatory Coherence Explain?" *Behavioral and Brain Sciences* 12 (1989), 475-76.

Glymour, C. et al. (1987), *Discovering Causal Structure: Artificial Intelligence, Philosophy of Science, and Statistical Modelling*. Orlando, FL: Academic Press.

Johnson-Laird, P.N. (1983), *Mental Models*. Cambridge: Harvard Univ. Press.

Kuhn, T.S. (1962), *The Structure of Scientific Revolutions*. Chicago: Univ. of Chicago Press (2nd ed. 1970).

______. (1977), *The Essential Tension*. Chicago: Univ. of Chicago Press.

Kulkarni, D., and Simon. H. (1988), "The Processes of Scientific Discovery: The Strategy of Experimentation". *Cognitive Science*, 12:139-175.

Lakatos, I. (1970), "Falsification and the Methodology of Scientific Research Programmes". In *Criticism and the Growth of Knowledge*, ed. I. Lakatos and A. Musgrave. Cambridge: Cambridge Univ. Press.

Lakoff, G. (1987), *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind*. Chicago: Univ. of Chicago Press.

Langley, P., Simon, H.A., Bradshaw, G.L., and Zytkow, J.M. (1987), *Scientific Discovery*. Cambridge: MIT Press.

Larkin, J.H., and Simon, H.A. (1987), "Why a Diagram is (Sometimes) Worth a Thousand Words". *Cognitive Science*, 11: 65-99.

Laudan, L. (1977), *Progress and Its Problems*. Berkeley: Univ. of California Press.

______. (1984), *Science and Values*. Berkeley: Univ. of California Press.

Neisser, U. (1976), *Cognition and Reality*. New York: Freeman.

Nersessian, N.J. (1984), *Faraday to Einstein: Constructing Meaning in Scientific Theories*. Dordrecht: Nijhoff.

_________. (1988), "Reasoning from Imagery and Analogy in Scientific Concept Formation". In *PSA 1988*, vol. 2, ed. A. Fine and J. Leplin, 41-47. East Lansing, MI: The Philosophy of Science Association.

_________. (1991), "How do Scientists Think? Capturing the Dynamics of Conceptual Change in Science". In *Cognitive Models of Science*, ed. R.N. Giere, Minnesota Studies in the Philosophy of Science, vol. 15. Minneapolis: Univ. of Minnesota Press.

Nowak, G. and Thagard, P. (1991), "Copernicus, Ptolemy, and Explanatory Coherence". In *Cognitive Models of Science*, ed. R.N. Giere, Minnesota Studies in the Philosophy of Science, vol. 15. Minneapolis: Univ. of Minnesota Press.

Richards, R. (1987), *Darwin and the Emergence of Evolutionary Theories of Mind and Behavior*. Chicago: Univ. of Chicago Press.

Shapin, S. (1982), "History of Science and its Sociological Reconstructions". *History of Science* 20:157-211.

Shrager, J. and Langley, P. (1990), *Computational Models of Discovery and Theory Formation*. Palo Alto, CA: Morgan Kaufmann Publishers, Inc.

Simon, H.A. (1978), "On the Forms of Mental Representation". In *Perception and Cognition: Issues in the Foundations of Psychology*, ed. C.W. Savage, 3-18, Minnesota Studies in the Philosophy of Science, vol. 9. Minneapolis: Univ. of Minnesota Press.

Suppe, F. (1989), *The Semantic Conception of Theories and Scientific Realism*. Urbana, IL: Univ. of Illinois Press.

Thagard, P. (1988), *Computational Philosophy of Science*. Cambridge: MIT Press.

______. (1989), "Explanatory Coherence". *Behavioral and Brain Sciences*. 12: 435-467.

______. (1991), *Conceptual Revolutions*. Princeton: Princeton Univ. Press.

van Fraassen, B.C. (1980), *The Scientific Image*. Oxford: Oxford Univ. Press.

__________. (1989), *Laws and Symmetry*. Oxford: Oxford Univ. Press.

https://doi.org/10.1086/psaprocbienmeetp.1990.2.193085 Published online by Cambridge University Press