Philosophy of Science

Pete Mandik, William Paterson College, Wayne, New Jersey, USA
William Bechtel, Washington University in St Louis, St Louis, Missouri, USA

CONTENTS

The logical structure of science
Challenges to Logical Positivism
The sociohistorical structure of science

Philosophy of science concerns the principles and processes of scientific explanation, including both processes of confirmation and of discovery.

00192001 Philosophy of science is primarily concerned to provide accounts of the principles and processes of scientific explanation. Early in the twentieth century, philosophers of science focused on the logical structure of scientific thought, whereas in the later part of the century logic was de-emphasized in favour of other frameworks for conceptualizing scientific reasoning and explanation, and an emphasis on historical and sociological factors that shape scientific thinking. While tracing through the landmarks of this history we note many points of contact between the philosophy of science and the cognitive sciences.

THE LOGICAL STRUCTURE OF SCIENCE

The Deductive-Nomological Model of Explanation and Hypothetico-Deductive Model of Theory Development

00192002 The appeal to logic to articulate the structure of scientific explanation and scientific reasoning was the hallmark of the Logical Positivists, a group of early twentieth-century philosophers and scientists, working initially in Eastern Europe, who sought to provide an explication of science that could explain its high epistemic status. They offered a model of explanation, the Deductive-Nomological (D-N) model, which holds that explaining a phenomenon involves deducing its occurrence from laws (Hempel, 1966). Something like this view is thousands of years old and may be discerned in the work of Aristotle. The D-N model was extremely influential in some areas of psychology earlier in the twentieth century. Many behaviourists, for example, sought to discover general laws of learning to characterize how various kinds of experiences (e.g. reinforcement) would change the behaviour of organisms.

Laws, which are central to the D-N model, are taken to specify general relations (as in Newton's law that force equals mass times acceleration ($f = ma$)). To apply these general relations to particular events, one must specify conditions holding at a previous time, which are usually called *initial conditions*. Recognizing that multiple laws and initial conditions may be involved in a given explanation, such explanations can then be represented in the following canonical form (where L designates a law, C an initial condition, and E the event to be explained): 00192003

L1, L2, L3, . . .
C1, C2, C3, . . .
Therefore E

Advocates of the D-N perspective generally assumed that the Cs and Es were sentences whose truth or falsity could be determined directly through observation. These *observation* sentences also provided the empirical support for the laws. Most advocates of a logical analysis of science disavowed interest in scientific discovery, restricting their focus to justification. Their characterization of the relationship between observations and laws make it clear why they saw little hope for a logic of discovery. (More recently, artificial intelligence (AI) researchers and philosophers sympathetic to AI have proposed discovery tools to employ within this framework: Langley *et al*., 1987; Thagard, 1988; Darden, 1991.) The challenge in discovery is to construct new laws to account for more or new phenomena. Since laws are to be general, not specific, they must go beyond any finite set of data, and thus there inevitably seems to be a major inductive leap to a new hypothesis. Once a hypothesis was put forward, the process of testing could begin by deducing its observational consequences in accord with the D-N model. This process of hypothesizing general laws and testing their deduced consequences was termed the Hypothetico-Deductive model of theory development.

00192004 Hempel (1966) illustrates H-D with the example of Semmelweis' work during the 1840s on childbed fever. Semmelweis observed cases of 'puerperal fever' or 'childbed fever' contracted by women who delivered children in his hospital. He noted that cases were especially frequent in groups of women where deliveries were handled by physicians instead of midwives. Semmelweis' key insight into the cause of childbed fever came when he observed that a physician came down with similar symptoms upon injuring himself with an instrument during an autopsy. Semmelweis hypothesized that 'cadaveric material' on the injurious instrument caused the disease, and that, similarly, the physicians associated with outbreaks of childbed fever had cadaveric material on their hands prior to delivering babies. Semmelweis tested this hypothesis by examining its implications. One implication of the hypothesis that cadaveric material is the cause of childbed fever is that its removal from the hands of physicians would result in a decrease in cases of childbed fever. When Semmelweis tested this implication by requiring that physicians wash their hands in chlorinated lime prior to examining patients (which he assumed would remove the cadaveric matter), he observed that groups of women examined by physicians who washed with chlorinated lime had lower incidents of childbed fever than groups of women examined by physicians who did not.

00192005 The example of Semmelweis' hypothesis and test conforms to the H-D model in the following way. Semmelweis' observations could be formulated as a series of observation statements, statements of particular states of affairs such as 'Jane Doe was exposed to cadaveric material and contracted childbed fever', 'Mary Smith was exposed to cadaveric material and contracted childbed fever', and so on. His hypothesis took the form of a law-like general statement, 'Any woman exposed to cadaveric material will contract childbed fever'. And, in accordance with D-N, the original observation statements may be deductively inferred from the statement of law. Thus the relation of observation statements to statements of law has a reciprocal structure as depicted by H-D and D-N. (See Figure 1.)

00192006 If the only support for a hypothesis were the observation sentences it was intended to explain, the reciprocal relationship between proposing hypotheses and explanation would be circular. Circularity is avoided, though, since a law, being a general statement, deductively entails not just the particular observation statements it was advanced to explain, but an indefinite number of others. The theory is confirmed by demonstrating the truth of some of these additional observation statements.

Intertheoretic Reduction

00192007 Recognizing that one might want to explain why laws held, the proponents also generalized this framework, allowing for the derivation of one or more sets of laws (each comprising a theory) from another set of laws (comprising another theory). The second set of laws will be more general ones from which, under specific boundary conditions, the first set of laws might be derived. (Thus, the boundary conditions replace the initial conditions in the above formalism.) Proponents also suggested that this approach might be extended to relations between laws in one science and those of a more basic science by providing bridge laws relating the vocabularies of the two sciences (by translating the terms of one into the terms of another), giving rise to the following schema:

Laws of the lower-level science
Bridge laws
Boundary conditions
―――――――――――――――
Laws of the higher-level science

These derivations are known as *reductions*; they figure prominently in discussions about the relation between psychology and neuroscience in which some theorists propose that the theories of psychology ought to reduce to those of neuroscience (Churchland, 1986). Analogously, reductions are posited, or at least hoped for, for any two adjacent levels, such as from biology to chemistry and from chemistry to physics.

00192008 One example of a successful reduction that conforms to the above account is the derivation of the Boyle–Charles Law of classical thermodynamics (specifications of the temperature and pressure relations in an ideal gas) from statistical mechanics. Boyle–Charles Law terms such as *temperature* and *pressure* are translated into the terms of statistical measures of the kinetic properties of molecules in a volume. Equating temperature with mean kinetic energy supplies one of the bridge laws enabling translations of the laws of Boyle–Charles into the laws of statistical mechanics. Boundary conditions include specifications of the kinds of molecules, the volume to which their motion is restricted, and the range of temperatures and pressures they are subject to. With translations of terms and boundary conditions in place, the Boyle–Charles Law is derivable from the laws of statistical mechanics.

00192009 Just as laws gain their empirical support from the true observation sentences that are derived from them, so reducing laws gain their support from the already confirmed laws that can be derived from them. Reduction, though, can also provide justification for reduced laws. Insofar as the reducing laws are more general, confirmation they receive in some domains can provide indirect support for other laws that can be derived from them.

CHALLENGES TO LOGICAL POSITIVISM

Popper's Critique of Confirmation

00192010 By making predictions which turn out true, the Logical Positivists thought we could justify laws. On this claim, however, they were criticized by Karl Popper (Popper, 1935/1959), who noted that such arguments had the invalid form of affirming the consequence:

If L were true, then prediction P would be true
P is true

L is true

This formalism is invalid since it is possible for both premises to be true, but the conclusion false. Instead, Popper argued that the only way evidence could bear upon laws was through the use of *modus tollens* arguments in which failed predictions could be used to falsify a purported law:

If L were true, then prediction P would be true
P is false

L is false

Accordingly, Popper emphasized that the method of science was a method of conjectures and refutations in which scientists proposed explanatory laws and then sought evidence showing that they were false. If a proposed law resisted all attempts at falsification, Popper would speak of it as corroborated, not as true or confirmed, recognizing that future evidence could always reveal it to be false.

Mechanisms instead of Laws

00192011 Although the D-N model seems to apply well to a number of scientific domains, especially in physics (for example, to explanations of phenomena that appeal to the laws describing ideal gases or principles of thermodynamics), it does not seem applicable to many domains in the life sciences. A major reason is that explanation does not tend to involve showing that a phenomenon follows a law (Cartwright, 1983; Giere, 1999). Instead, explanation often involves identifying and describing the mechanism that generates the phenomenon (Wimsatt, 1972; Machamer *et al*., 2000). A key component to the idea of mechanism is that of a set of processes that generate a phenomenon, with these processes being performed by different parts of a system. Thus, an explanation consists of functionally decomposing the process of producing the phenomenon into a set of different component processes and localizing these component processes in actual physical parts of a system (Bechtel and Richardson, 1993). (Often actually identifying the component physically is not possible, and researchers settle for indirect evidence that such a component exists.) For example, to explain basic physiological processes such as cellular respiration, biochemists had to identify a set of activities (oxidation of substrates, electron transfer, and phosphorylation of ADP) and determine the components of the cell responsible for them (enzymes, cofactors, and membranes with restricted permeability).

An emphasis on mechanisms does not eliminate 00192012
appeal to laws – sometimes key relations between parts of a system (e.g. between a substrate and an enzyme) are expressed in laws. But laws play a subsidiary role. The emphasis is on differentiating the operations performed in the system, linking them with physical parts of the system, and then showing how the component parts and processes interact with each other to produce the phenomenon. Some of the individual interactions can be stated in laws, but the specification of the particular components involved and the intricacy of their interactions is generally far too specific to render into laws. Quite often such explanations are presented in diagrams (these are especially useful when the component processes are organized nonlinearly with multiple interactions and feedback loops), not linguistically (although diagrams are typically accompanied by linguistic commentary).

Many of the endeavours of the cognitive sciences 00192013
can be interpreted as advancing mechanisms. Grammars, for example, are often presented as mechanisms for generating sentences. Psycholinguists who investigate the psychological reality of particular grammars are investigating whether they are actually implemented in language users. Researchers in AI, especially those creating programs to account for human performance, are decomposing an activity into component operations. Implementing the program on a computer

provides an existence proof that the hypothesized set of operations is sufficient to generate the phenomenon in question. Behavioural experiments using such tools as reaction times and error analysis are required to demonstrate that the processes actually figure in human information processing.

00192014 When the explanatory vehicle is assumed to be a description of a mechanism, scientific inquiry is not restricted to producing data to be subsumed under a hypothesis or to being used in testing a prediction derived from a hypothesis. Early in the process of inquiry, researchers are simply trying to figure out what are the processes that contribute to a particular effect. One strategy is to show that processes are actually separable in the system by showing that one process can be impaired while the other is retained, or that an experimental manipulation produces a crossover interaction between measures of the two processes. When these manipulations are performed after a mechanism has been proposed, they serve the more traditional role of testing the proposed explanation. In such ways, empirical inquiry contributes both to discovery of models of mechanism and to testing them.

Alternative Conceptions of Reduction

00192015 The application of the model of mechanistic explanation in the previous paragraphs focused on the traditional disciplines of cognitive science: cognitive psychology, linguistics, and AI. But a very natural place to employ this framework is to relations between more traditional cognitive explanations and neuroscientific ones. Although interest in the neural realization of cognitive mechanisms did not play a critical role in cognitive science until recently, this was largely due to the paucity of techniques that could link cognitive and neural investigations (Bechtel *et al.*, 1998). But the emergence of cognitive neuroscience as a major area of scientific collaboration in the 1990s reveals that cognitive and neural modes of investigation can be invoked together. Indeed, neuroimaging experiments, whether with positron emission tomography (PET), functional magnetic resonance imaging (fMRI), magnetoencephalography (MEG), or (ERP), require use of cognitive tasks and measures along with measures of neural activity. In terms of the model of mechanistic explanation, what these tools are providing is localization of the functions decomposed in a cognitive analysis of the task (e.g. by showing where various attentional processes are realized in the brain). Like more purely behavioural research, moreover, such studies can not only serve to confirm a cognitive decomposition, but also play a heuristic role in determining the functional decomposition itself (Bechtel *et al.*, 2000).

00192016 Such connections between cognition and neuroscience are commonly construed as reductionistic. But since one is not starting with laws at the cognitive level, and deriving them from laws of neuroscience, such research does not fit the theory reduction model (above). Rather, relating a functional decomposition to a structural localization provides an alternative conception of reduction, one much closer to Darden and Maull's (1977) conception of an interfield theory. Such an account may be much closer to actual scientific practice. Moreover, it does not raise the spectre of either explaining away the higher-level approach (with a successful reduction) or eliminating it (if reduction fails). Rather, it weaves the two approaches closely together; functional decomposition and structural localization are contributors to the common inquiry of understanding the mechanism. Each can provide not only support for the other, but heuristic guidance to the other, and the resulting explanatory account is an integrated one.

00192017 There is a further way in which reduction emerges in the context of a mechanistic model of explanation. Once a system has been decomposed into component functions and these localized within the system, a new explanatory task arises – explaining how each component function is performed. Developing this explanation requires repeating the process: decomposing the component process into its component processes, and localizing these within the subsystem where the process from the first decomposition was localized. Recognizing that this can be done successively, the model of mechanistic explanation provides a framework for reduction through multiple levels (Wimsatt, 1976). There is an important point to recognize about this multilevel conception of reduction – at each level, a different phenomenon is being explained. Moreover, at each level, the explanation involves not only the contributing components, but the interactions that are specified in the account at that level. A process such as seeing is reductively explained in terms of the contributions of the different components of the visual system and their interactions, while it is the activity of a given component which is explained by going into it and identifying its subcomponents and their interaction.

THE SOCIOHISTORICAL STRUCTURE OF SCIENCE

Paradigms and Revolutions

00192018 While the Logical Positivists were themselves very interested in the science of their time, their account was grounded primarily in logic, not in the details of scientific practice. (A consequence of this is that they viewed it as a normative model characterizing any possible science.) Kuhn's work drew philosopher's attention (as well as that of historians and sociologists of science) to the specific details of the process of scientific research. Kuhn (1996) challenged the views that science employed a set of general methods that remained constant over time and accumulated a body of truths. In their stead, Kuhn suggested that scientific approaches vary so significantly at different eras of a discipline that the findings and theories at one time cannot be meaningfully related to the findings and theories of other times. Instead of viewing the historical progression of science as the progressive accumulation of truths, Kuhn offered a cyclic model of the stages of scientific activity through history. The cycle involves the five stages of (1) immature science, (2) normal mature science, (3) crisis science, (4) revolutionary science, and (5) resolutions, which is followed by a return to normal science (see Figure 2).

00192019 The key notion in understanding Kuhnian philosophy of science is the notion of a *paradigm*. The five stages of the Kuhnian cycle may be unpacked in terms of this notion, since normal science is paradigm-based science, and four of the stages are understood by way of contrast with normal science. For Kuhn, paradigms are 'Universally recognized scientific achievements that for a time provide model problems and solutions to a community of practitioners' (1996, p. x). Further, paradigms serve for a time implicitly to define the legitimate problems and methods of a research field for succeeding generations of practitioners. They do so in virtue of two essential characteristics (p. 10): first, 'Their achievement was sufficiently unprecedented to attract an enduring group of adherents away from competing modes of scientific activity.' Second, their achievement 'was sufficiently open-ended to leave all sorts of problems for the redefined group of practitioners to resolve'. Examples of paradigms include Ptolemaic astronomy, Copernican astronomy, Aristotelian dynamics, Newtonian dynamics, corpuscular optics, and wave optics.

Normal science, then, just is paradigm-based science. Immature science is science studying a domain recognizably the same as that studied by paradigm-based successors, but without the utilization of any paradigms, as in the cases of optics prior to Newton and of electrical research in the first half of the eighteenth century (pp. 12–14). A stage of crisis emerges for a science when the requisite consensus regarding the applicability of a paradigm begins to unravel. Patterns of problem-solving in normal science give way to novel approaches, and where new consensus emerges, a stage of resolution of a new paradigm returns the cycle to normal science. 00192020

One of Kuhn's views that garners much attention in the philosophy of science is his claim that different paradigms are incommensurable and thus choice of one over another cannot be subject to rational procedures. One way of thinking of incommensurability of paradigms is by thinking of paradigms as involving different languages, the terms of which cannot be translated into each other. For example, the term 'space' as used within a Newtonian physical paradigm cannot be translated as the term 'space' as used within an Einsteinian physical paradigm, and vice versa. Einsteinians differ from Newtonians in holding that space is curved by mass, for example. The meaning of 'space' depends on the theory it is embedded in; where theories diverge, so do the meaning of their terms. Part of Kuhn's argument that paradigms are not open to rational choice hinges on the notion that observation statements cannot serve as neutral points of arbitration since there is no theory-neutral observation language. In other words, observation is theory-laden: how one perceives the world depends on the theory with which one conceives the world. For instance, a newborn baby, if shown a cathode-ray tube, would not see it *as* a cathode-ray tube because the baby would understand an insufficient amount of theory to know what cathode-ray tubes are. 00192021

Kuhn's thesis that paradigms are incommensurable leads to another of his theses, namely, that the history of a scientific discipline is noncumulative. Noncumulativity follows from incommensurability in the following way. Since the language of one paradigm cannot be translated into the language of another, the statements held to be true within one paradigm cannot be expressed, let alone judged to be true, within another. The theory-ladenness of perception guarantees that not only can items of theory not be accumulated across paradigms, but neither can observations. Since not even observations are retained from one 00192022 paradigm to the next, the history of a discipline cannot be viewed as the accumulation of truths or as progressing towards a truer account of the world. This challenges the traditional view of science as an intellectual that progresses over time. The history of science, as viewed through the Kuhnian lens, is of a series of paradigms and revolutions, none bearing any rational relation to any other.

Challenges to Kuhn

00192023 Kuhn's hypothesis of incommensurability has been challenged on several fronts. One denies that the meanings of theoretical terms are determined wholly by factors internal to a paradigm, but holds instead that they may be determined, at least in part, by causal relations between the term and items in the external world. Putnam (1975) suggests that the meaning of certain scientific terms involves causal relations between the terms and things in the world that they denote. For instance, part of the meaning of water is the substance H_2O that was present when the term 'water' was first brought into use to denote that substance. A causal chain leads from current uses of 'water' to the initial dubbing of H_2O as water. These causal chains remain constant regardless of a scientist's theory. Thus, water discourse need not be incommensurable between adherents of divergent theories about water. Appealing to causal relations to fix the content of representations is mirrored in much cognitive scientific practice, especially in the cognitive neurosciences where representations are thought of as carrying information about their environmental causes.

00192024 Another challenge to Kuhnian incommensurability arises from theorists who propose that the mind is modular. Fodor (1984), for example, argues that many perceptual processes are modular in the sense of being 'informationally encapsulated' so that their outputs are immune from influence by theoretical and other acquired beliefs. Fodor, therefore, contends that observational reports can be treated as univocal even when theorists hold different theories.

00192025 Though many of the issues raised by Kuhn continue to be debated, the Kuhnian spirit is a pervasive force in post-Positivist philosophy of science. It has, for example, led philosophers to focus much more on the diachronic nature of science as well as the actual research processes of science (see, for example, Lakatos, 1970; Laudan, 1977) or to apply Bayes Theorem to model rational theory choice (Howson and Urbach, 1993; Mayo, 1996). Nonetheless, no consensus has emerged in contemporary philosophy of science that is on a par with the status once accorded the Positivists. For cognitive scientists, one of the attractive features of current philosophy of science is the increasing effort to employ ideas from cognitive science in the attempt to understand science (Giere, 1992; Thagard, 1998).

References

Bechtel W, Abrahamsen A and Graham G (1998) The life of cognitive science. In: Bechtel W and Graham G (eds) *A Companion to Cognitive Science*, pp. 1–104. Oxford, UK: Blackwell.

Bechtel W, Mandik P and Mundale J (2000) Philosophy meets the neurosciences. In: Bechtel W, Mandik P, Mundale J and Stufflebeam RS (eds) *Philosophy and the Neurosciences: A Reader*. Oxford, UK: Blackwell.

Bechtel W and Richardson RC (1993) *Discovering Complexity: Decomposition and Localization as Scientific Research Strategies*. Princeton, NJ: Princeton University Press.

Cartwright N (1983) *How the Laws of Physics Lie*. Oxford, UK: Clarendon Press.

Churchland PS (1986) *Neurophilosophy*. Cambridge, MA: MIT Press.

Darden L (1991) *Theory Change in Science: Strategies from Mendelian Genetics*. Oxford, UK: Oxford University Press.

Darden L and Maull N (1977) Interfield theories. *Philosophy of Science* **43**: 44–64.

Fodor J (1984) Observation reconsidered. *Philosophy of Science* **51**: 23–43.

Giere R (1992) *Cognitive Models of Science*. Minneapolis, MN: University of Minnesota Press.

Giere R (1999) *Science without Laws*. Chicago: University of Chicago Press.

Hempel CG (1966) *Philosophy of Natural Science*. Englewood Cliffs, NJ: Prentice-Hall.

Howson C and Urbach P (1993) *Scientific Reasoning: The Bayesian Approach*. La Salle, IL: Open Court.

Kuhn TS (1996) *The Structure of Scientific Revolutions,* 3rd edn. Chicago: University of Chicago Press.

Lakatos I (1970) Falsification and the methodology of scientific research programmes. In: Lakatos I and Musgrave A (eds) *Criticism and the Growth of Knowledge*. Cambridge, UK: Cambridge University Press

Langley PS, Simon HA, Bradshaw GL and Zytkow JM (1987) *Scientific Discovery: Computational Explorations of the Creative Process*. Cambridge, MA: MIT Press.

Laudan L (1977) *Science and Relativism*. Berkeley, CA: University of California Press.

Machamer P, Darden L and Craver CF (2000) Thinking about mechanisms. *Philosophy of Science* **67**: 1–25.

Mayo D (1996) *Error and the Growth of Experimental Knowledge*. Chicago: University of Chicago Press.

Popper K (1935/1959) *The Logic of Discovery*. London: Hutchinson.

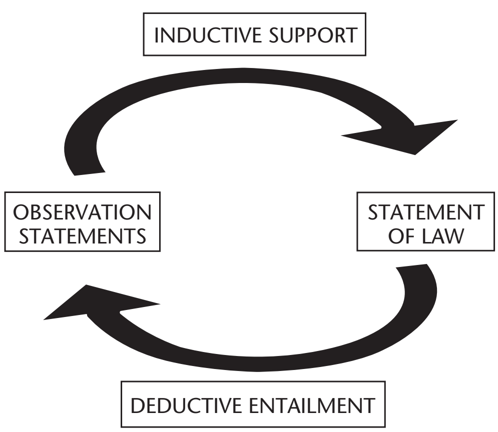
Putnam H (1975) The Meaning of 'Meaning'. In: *Mind, Language, and Reality: Philosophical Papers of Hilary Putnam*. Cambridge, UK: Cambridge University Press.
Thagard P (1988) *Computational Philosophy of Science*. Cambridge, MA: MIT Press.
Wimsatt WC (1972) Complexity and organization. In: Schaffner KF and Cohen RS (eds) *PSA* 1972, pp. 67–86. Dordrecht: Reidel.
Wimsatt WC (1976) Reduction, levels of organization, and the mind–body problem. In: Globus G, Maxwell G and Savodnik I (eds) *Consciousness and the Brain: A Scientific and Philosophical Inquiry*. New York: Plenum Press.

Further Reading

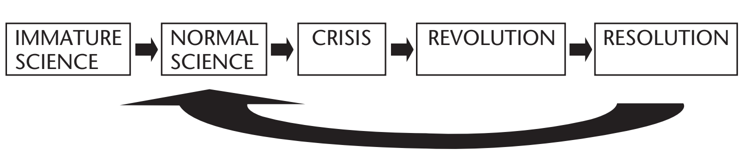
Bechtel W (1988) *Philosophy of Science*. Hillsdale, NJ: Lawrence Erlbaum.
Glymour CN (1980) *Theory and Evidence*. Princeton, NJ: Princeton University Press.
Hacking I (1983) *Representing and Intervening*. Cambridge, UK: Cambridge University Press.
Laudan L (1990) *Science and Relativism: Some Key Controversies in the Philosophy of Science*. Chicago: University of Chicago Press.
Simon HS (1996) *The Sciences of the Artificial*, 3rd edn. Cambridge, MA: MIT Press.
Thagard P (1992) *Conceptual Revolutions*. Princeton, NJ: Princeton University Press.

Keywords: (Check)

Explanation; mechanism; reduction; theory change; scientific discovery



00192f001 **Figure 1.** The reciprocal relationship between hypothesizing laws from observations and deriving observations from laws



00192f002 **Figure 2.** The stages in Kuhn's account of scientific change

ECSaq192
Queries for Macmillan, ECS paper no. 192

Title: Philosophy of Science
Author: Bechtel

Can you supply a small glossary defining complex terms for the layperson?

Displayed syllogisms (e.g. L1 etc., C1 etc., Therefore E): in each case, the conclusion (e.g. Therefore E) is preceded by a box - is this correct? Should it in fact be the symbol "therefore" (three dots), and was incorrectly rendered in the computer print-out?

Section "Mechanisms instead of Laws", 1st para: we'll need to give ADP in full.

Section "Alternative Conceptions of Reduction": similarly, we'll need PET, fMRI, MEG, ERP in full: are the first three "positron emission tomography", (functional) magnetic resonance imaging, and magnetoencephalography? (I just can't remember what ERP is, I'm afraid . . .)

Section "Paradigms and Revolutions", end of last para: "traditional view of science as an intellectual [? word missing ?] that progresses over time" - ?

End of main text: no reference is given for Thagard 1998.

Giere 1992 ref: is the place of publication Minneapolis, MN?

Wimsatt 1972 ref: what is PSA (title)? Proceedings of - ?