# A Problem-Solving Account of Scientific Explanation

Gary L. Hardcastle
Department of Philosophy, 0126
Virginia Tech
Blacksburg, VA 24061-0126
garyh@vt.edu http://truth.phil.vt.edu/

# A Problem-Solving Account of Scientific Explanation

Abstract

An account of scientific explanation is presented according to which (1) scientific explanation consists in solving "insight" problems (Metcalfe and Wiebe 1984) and (2) understanding is the result of solving such problems. The theory is pragmatic; it draws upon van Fraassen's (1977, 1980) insights, avoids the objections to pragmatic accounts offered by Kitcher and Salmon (1987), and relates scientific explanation directly to understanding. The theory also accommodates cases of explanatory asymmetry and intuitively legitimate rejections of explanation requests.

## 1. The Problem of Scientific Explanation

The explanans of any genuine scientific explanation bear a relation of explanatory relevance to that explanation's explanandum. The central philosophical task associated with scientific explanation is the description of this relation, and competing philosophical theories of scientific explanation can often be understood as singling out some relation as "the" relation of explanatory relevance. However, another way to respond to this philosophical task is to hold that there is no single correct relation of explanatory relevance, but that what counts as the appropriate relation varies with facts about the people giving, and getting, the explanation. This response characterizes pragmatic theories of scientific explanation, for on such views explanans can bear on explananda in a genuine scientific explanation in any number of ways, depending on the people involved (cf. Hempel 1965, pp. 425-433 and Humphreys 1989, pp. 126-127; see also Section 4 below).

In this section I examine van Fraassen's (1977, 1980) pragmatic account of explanation (for other accounts see Achinstein (1983, 1986) and Churchland (1989)), and the objection leveled against it by Kitcher and Salmon (1987). In Section 2 I describe two approaches to scientific explanation, and reasons for pursuing one over the other. In Section 3 a sketch of some features of understanding and some experimental studies of "insight problems" motivate the problem-solving account of explanation described there. The problem-solving account of explanation is a pragmatic account in the spirit of van Fraassen's; its central tenants are that (1) scientific explanation consists in solving "insight" problems, and (2) understanding is the result of solving such problems. The final section addresses some objections and touts the virtues of the problem-solving account -- it accommodates both explanatory asymmetries and legitimate rejections of explanation requests, is immune to Kitcher and Salmon's objection to pragmatic accounts, and connects scientific explanation and understanding.[i]

Van Fraassen's account is cast within an *erotetic* framework, in which scientific explanations are presented as answers to explanation-seeking why-questions. Pragmatic and erotetic accounts of explanation often go hand in hand (as Humphreys (1989, p. 133) has noted), but they need not. In presenting van Fraassen's account below I will use some familiar erotetic machinery, but it is not part of my view that all explanations are, or can be framed as, answers to explanation-seeking why-questions; indeed, it seems clear that some cannot (see Salmon 1989, 137-138)

Van Fraassen's account specifies as the formal structure of an explanation-seeking why-question Q an ordered triple $<P_K, X, R>$ offered against background knowledge K, in which $P_K$ is a proposition (the *topic* of Q); X is a *contrast* class of propositions $\{P_1, P_2, P_3, \ldots P_K, \ldots\}$; and R is a *relevance relation*. On this account, for example, the interrogative 'Why did the Titanic sink?' is typically elliptical for the question $<P_K, X, R>$, where $P_K$ is the proposition that the Titanic sunk, X is {The Titanic sunk, The Titanic did not sink} and R is the relation of causation. The same interrogative may of course represent a different question, say, one with a different contrast class or relevance relation (see van Fraassen 1980, p. 142; Salmon 1989, pp. 139-40).

In general, to pose Q is to presuppose that $P_K$ is true, that $P_K$ is the only true member of X, and that there is some true proposition A which bears R to $<P_K, X>$. It is appropriate to raise Q only in contexts where these three presuppositions are satisfied. Specifically, Q *arises* in K iff K entails

(a) $P_K$ is true and

(b) no member of X other than $P_K$ is true,

and K does *not* entail that

(c) no true A stands in R to $\langle P_K, X\rangle$.

Questions posed in contexts where they do not arise are, technically speaking, mistaken. Mistaken questions admit corrective answers, which consist in pointing out either that K does not entail (a)-(b) or that it does entail (c). A corrective answer may lead to a reformulated explanation-seeking why-question. If asked why the Titanic struck another ship (rather than nothing), for example, the appropriate (corrective) answer is to point out that the Titanic did *not* strike another ship (i.e., the topic is false) and add perhaps that it *did* strike an iceberg, suggesting the question, “Why did the Titanic strike an iceberg (rather than nothing)?” In such fashion van Fraassen’s account manages one task of a theory of explanation, namely, to “account for legitimate *rejections* of explanatory requests” (van Fraassen 1980, 146). A question is legitimately rejected in a context K just if it does not arise on K. Unless otherwise noted, the explanation-seeking why-questions discussed below will be presumed to arise.

A direct *answer* to Q is a proposition of the form “$P_K$ in contrast to the rest of X, because A”, where

(I) A is true,

(ii) $P_K$ is true,

(iii) Each $P_j$ in X is false if $j \neq k$, and

(iv) A bears R to $\langle P_K, X\rangle$ (van Fraassen 1980, pp. 144-145).

A is the core of the direct answer, for in bearing R to the one true member of X, A provides what was, intuitively, requested by Q -- a "singling-out" of $P_K$ among the other members of X by virtue of R.

Some answers to an explanation-seeking why-question are better than others; some, in van Fraassen's words, are more "telling." Van Fraassen (1980, §4.4, esp. pp. 146-147) suggests that an answer to Q with core A is telling to the extent that:

(1) $Pr(A|K)$ is high,

(2) A favors $P_K$, where to favor $P_K$ is to maximize the difference between $Pr(P_K|A \& K(Q))$ (where K(Q) is a subset of the background knowledge K), and the probabilities of the other propositions in X similarly conditionalized, and

(3) There is no rival to A which is more probable in light of K, which favors $P_K$ more than A, or which renders A statistically irrelevant to $P_K$.

The probabilities in (1)-(3) are presumably to be construed subjectively (see Salmon 1989, p. 145).

Following Kitcher and Salmon's (1987, pp. 319-320) terminology, call maximally telling answers perfect answers, questions well-founded to the extent that they admit of telling answers, and questions which admit of perfect answers maximally well-founded. Kitcher and Salmon (1987, p. 321) show that for any true $P_K$ and true A (not necessarily distinct) in K there is a Q with $P_K$ as its topic such that A is an essential part of a perfect answer to Q. Let A and $P_K$ be true members of K and X a set of propositions the only true member of which is $P_K$. Let Z be the disjunction of the members of X other than $P_K$, and let K entail ~Z. Let R be the relation that holds between a proposition B and $P_K$ just if B entails $P_K$. Then

B: A • (A --> $P_K$) • ~Z

is the core of a perfect answer to Q, for Pr(B|K) = 1, B entails $P_K$ and ~Z (and so, in van Fraassen's (1980, p. 147) words, "receives... the highest marks for favoring the topic,") and no other answer to Q can favor $P_K$ more strongly or render B irrelevant. What is to be noted here is that A, an essential part of the perfect answer to Q, can be any true member of K at all, no matter how intuitively irrelevant to $P_K$.

Kitcher and Salmon offer the following illustration of this point (1987, p. 322). Adapting somewhat from their example, let

$P_K$: JFK died on 11/22/63,

X: {JFK died on 1/1/63, ... JFK died on 12/31/63, JFK survived 1963}, and

R: Logical consequence, i.e., R holds between A and B just if B is a logical consequence of A.

Let A be a true description of the relative position of heavenly bodies at the time of JFK's birth. Then the question, "Why did JFK die on 11/22/63 (rather than on another date in 1963 or not at all that year)?" has as a perfect answer

B: A • (If A, then JFK died on 11/22/63) • JFK did not die on 1/1/63 • ... • JFK did not die on 11/21/63 • JFK did not die on 11/23/63 • ... • JFK did not survive 1963 (Kitcher and Salmon 1987, 232).

an essential part of which is the true description of the relative position of various heavenly bodies at the time of JFK's birth. Obviously, more far-fetched examples can easily be conjured.

Kitcher and Salmon view this result "as a reductio of van Fraassen's account of explanation" (1987, p. 319), and they claim that repairing the account requires adding as a fourth presupposition of Q,

(d) R is a relevance relation (1987, 322)

which must be satisfied if Q is to not be rejected.[ii] (d) is intended to permit relations which correspond to objective relevance and prohibit those on which A is 'relevant' to $P_K$, as Salmon puts it, "only in the Pickwickian sense" (1989, p. 141). This is, of course, no friendly amendment to van Fraassen's philosophy of science; the addition of (d) would severely compromise his constructive empiricism insofar as it provides a way for explanatory power to objectively distinguish empirically equivalent theories (Kitcher and Salmon 1987, p. 329).

The problem-solving account of explanation offered here suggests a different response to the kinds of examples Kitcher and Salmon offer. On this view, explanation is the solving of a particular kind of problem and understanding is a psychological consequence of solving problems of this sort. "Explanations" like the one in the JFK example fail as explanations not because the relevance relations they employ are in fact not relevant but because they present problems that are, in a sense to be discussed below, trivially soluble. To borrow the expression favored by Salmon, these examples pose problems that are "solved" only in a Pickwickian sense; this is why they strike us as pseudo-explanations which contribute not at all to our understanding. Indeed, on the problem-solving account the desire to debar explanations which make use of bad science comes out as misguided, for when explanation is understood in terms of its putative telos, understanding, and understanding is understood as the psychological phenomenon it is, we can easily see how and in what sense theories that are false, implausible, or discredited might nonetheless provide explanations. As I shall argue in Section 4, what is called for instead of Kitcher and Salmon's (d) is a condition requiring of explanation-seeking why-questions that the problems they set be non-trivial.

The examples Kitcher and Salmon offer leave philosophers of science with a particular task vis a vis scientific explanation. On Kitcher and Salmon's view, the philosophical task of identifying one particular relation (or a small class of relations) as the right one in all contexts for purposes of scientific explanation finds its place within the more general philosophical task of identifying those relations which are not explanatorily relevant in any context. The notion that there is one genuine relevance relation (or a small class of such relations) operative in scientific explanations independent of the context Kitcher and Salmon dub "uniformitarianism." Uniformitarianism shares with "relativism" -- the view that "the set of genuine relevance relations [is] a function of the branch of science and of the stage of its development" (Kitcher and Salmon 1987, p. 325) -- the conviction that "there are some relations that are not genuine relevance relations at any historical stage of any science" (1987, p. 326). Identifying these pseudorelevance relations is thus the initial philosophical task.

Pragmatism, construed as the view that for any relation there is some context in which it is explanatorily relevant, does not appear in Kitcher and Salmon's catalog of philosophical positions.[iii] In the next section, I will examine how we should approach the task set by Kitcher and Salmon. A consequence of that study is the resuscitation of a thoroughly pragmatic account of scientific explanation.

## 2. Historical Versus Understanding-Based Approaches to the Problem of Explanation

In their discussion of the JFK example, Kitcher and Salmon dismiss astral influence qua relevance relation because it is at odds with our "present lights" concerning how the world works (1987, p. 322). Assessing candidate relevance relations (and the explanations that depend on them) in light of current science is a common strategy in philosophical discussions of explanation, one which gives force to the examples of flagpoles (and their shadows), barometers (and storms), and bald school board members. It is in the context of current science, after all, that the height of flagpoles explains their shadows (and not vice-versa), storms explain barometers (and not vice-versa), and school board membership is not (typically) an explanation of baldness.

This use of current science is part of a far more general strategy of measuring theories of explanation against the canon of recognized scientific achievement -- the body of episodes, theories, and experiments that qualify as exemplary science (for discussion of this canon see Laudan 1990, p.47, p.53; for its role in testing theories of explanation see Friedman 1974, p.13). A theory of explanation is required to account for the instances in this canon -- the more, and the wider the range, the better. For the purposes of this paper, call this approach to the problem of scientific explanation the "historical approach." With respect to the task of identifying relations which never figure in genuine explanations, the historical approach amounts to surveying the canon to demonstrate the absence of various relations, such as astral influence. Such absence is taken to indicate that astral influence is not a genuine relevance relation.

What other approach could there be to the problem of scientific explanation, but to formulate a theory of explanation and test it against cases antecedently recognized as extremely good (or extremely bad) scientific explanations? In a well-known paper, Michael Friedman describes an alternative, an approach which uses our notion of scientific understanding to "judge the adequacy of philosophical theories of explanation" (1973, p.6). From a demonstration that various theories of explanation do not "connect" explanation and understanding, Friedman proposes to "extract some general properties a concept of scientific understanding ought to have" and to build on this "an account of scientific explanation that possesses these desirable properties" (1974, p.6; Salmon (1998, C. 5) has lately emphasized the significance of understanding also). The "unificationist" theory of scientific explanation at which Friedman arrives in his paper has been the subject of much elaboration and discussion (see esp. Kitcher 1976, 1981, 1989, and 1993), but my interest here is in Friedman's approach to the problem of scientific explanation, one I will call "understanding-based." [iv]

These two approaches to scientific explanation are different, but they are not at odds. In one case we begin with detailed scrutiny of scientific practice past and present, good and bad, intending to emerge from this study with a compact account of explanation which "saves" these instances and, one hopes, instances yet to be examined. In the other we begin with a theory of understanding -- its features, function, and relation to other cognitive states -- and build from this a theory of what practice or practices promote understanding. _Ideally_, these two approaches, pursued separately, mesh in the end, i.e., the practices that lie behind and promote understanding turn out to form a compact description of the instances in the historical canon of scientific explanation, and the theory of explanation proffered by the historical approach describes a practice which promotes understanding on our more sophisticated account of it. In the _actual_ world it would be sufficient if the two approaches fit together well enough such that from them we have a single theory of explanation, aligned with a theory of understanding on one hand and an interpretation of the canonical instances of scientific explanation on the other. But the prospect of a happy ending should not disguise the fact that the problem of scientific explanation has at least two very different starting points.

Other philosophers have recognized that understanding provides a path to a theory of scientific explanation. Paul Churchland, for example, proposes that explanatory understanding is continuous with perception, and “consists in the activation of a particular prototype vector [in the hidden layers of nodes] in a well-trained [PDP] network” (1989, 210). On the basis of this view of understanding he proposes a “prototype activation” account of scientific explanation: explanation is the activation of a prototype vector constitutive of understanding in a network of parallel distributed processors like we might well find in the brain.

What is missing from Churchland’s account, and from Friedman’s as well, is a plausible _theory_ of human understanding. Churchland’s claim that understanding just _is_ the activation of a prototype vector is puzzling at best, for Churchland offers no reason _why_ understanding should be so construed -- no empirical support for the claim that understanding and prototype activation are correlated, and no argument for why the activation of a prototype vector (rather than another of the trillions of vectors that might be activated) should be associated with understanding. Friedman, on the other hand, never delivers the general properties of scientific understanding he promises; he settles instead for "three desirable properties that a theory of explanation should have" (1974, p. 13; emphasis added). The third of these properties is that the account of explanation "somehow connect explanation and understanding -- it should tell us what kind of understanding scientific explanations provide and how they provide it" (1974, p. 14). But in Friedman's paper, pretheoretic intuitions about when we have understanding and when we don't take the place of a theory of understanding against which we might measure an account of explanation.[v]

Short-circuited attempts at an understanding-based approach to explanation should not warn us off the approach itself. On the contrary, the understanding-based approach is a relatively untried and promising one, provided it can make use of a plausible theory of understanding. In the following section, I sketch a theory of understanding and present an account of scientific explanation inspired by it.

## 3. A Problem-Solving Account of Scientific Explanation

Human understanding is a rich and varied psychological phenomenon, to say the least (for a taxonomy of types of understanding, see Salmon (1998)). We can get some handle on the notion by considering some features of our "everyday," or pretheoretic, concept of understanding.

First, understanding on our usual, pretheoretic, notion is a cognitive state, rather than a cognitive capacity or process. We talk of "coming to understand," "gaining an understanding," "having (and loosing) an understanding," and of "almost (but not quite) understanding"; implicit in all of this talk is the idea that to understand X is to stand in some relation to X, i.e. some state, rather than to go through a series of states (i.e. a process) or to be able to bring about some state (a capacity). This is not to claim that understanding cannot be (or is not often) the result of a process, nor that there are not capacities for understanding; either of these claims is consistent with regarding understanding as a state. Nor is it quite to claim that processes or capacities may not themselves be ultimately characterized as states; the point is simply that understanding is typically regarded as a state, not a process or capacity.

Within the notion of cognitive state there is considerable leeway. Understanding may, for example, presumably be alternately standing or occurrent, in the way that a belief or desire may be standing sometimes, occurrent others. Perhaps understanding in the usual sense is semantically evaluable in the way that beliefs and some other cognitive states are, but this is less clear. Most theories of explanation bar "explanations" in which the explanans (or, in the erotetic framework, the answer) are false, presumably on the grounds that such "false explanations" would provide false understanding, and that notion is incoherent. To understand at all is to understand on the basis of truth, or, at least, not on the basis of falsehoods. Accordingly, understanding at least involves semantically evaluable (specifically, true) states, if it is not one itself. The account of explanation offered here does not quite follow this line, though, for reasons discussed below.

Second, understanding in its pretheoretical guise is discrete: one either understands X or one does not. *Prima facie* this is false, for we speak of *gradual* understanding and *degrees* of understanding, and of understanding something more, or better than, someone else. Indeed, several philosophers have presumed that understanding is not discrete. On Churchland's account, for example, understanding is the activation of a prototype vector and activated vectors can be nearer or farther from a prototype; thus one might understand X in degrees. And Kitcher (1985, p. 633) suggests classifying explanations as more or less complete on the grounds that "there are degrees of understanding, and of *un*understanding"; an ideally complete explanation eliminates all our ununderstanding.

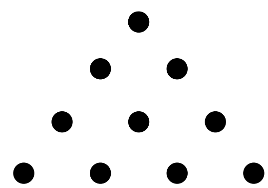
However, instances of understanding which initially appear non-discrete typically break down, upon analysis, into discrete components. Thus the gradually increased understanding one might have of a combustion engine, for example, is on analysis the accumulation of a set of *discrete* understandings -- of the fuel system, cooling system, exhaust system, etc., for example. This is probably too coarse even; understanding an electrical system, for example, is accumulating discrete understandings of other, simpler, phenomena. Our talk about understanding reflects its discrete character too, for we talk of understanding in a "flash" or an "instant", and we distinguish between those who do and those who do not understand something. The persistent fact that putatively gradual understandings can typically be analyzed into discrete components speaks against giving weight to talk of degrees of understanding, and suggests instead that we construe understanding X *more* (or *less*) as understanding, discretely, more (or fewer) of X's *components*.

Finally, understanding is, phenomenologically, both surprising and pleasant. These are not obvious features of our ordinary notion of understanding, and some care must be taken to disentangle them from some obvious truths about understanding. To begin with, this feature is not implied by understanding's status as a discrete state, for an arrival at a discrete state need not be phenomenologically sudden or surprising. A train at a platform is in a discrete state, but its arrival there is no surprise to those who watched it come down the track and into the station. However, understanding _is_ surprising; it is like a train that suddenly appears at the station. It is not, for that, necessarily unexpected. One can stare at a proof for hours, knowing that eventually its workings will stand forth clearly. Still, the grasp of the proof comes, when it does come, suddenly and despite this expectation. This is a prevalent feature of understanding.

We must also distinguish the effect of understanding _that X_ from the effect of simply understanding. Often the two are similar, as when for example when one grasps that the coins inherited from a deceased relative are untarnished because they are gold. There is in this case some pleasure had as a result of understanding why these coins have not tarnished, and much more from the realization of an impending financial windfall. In cases where we are indifferent or adverse to that which we come to understand, though, the pleasurable aspect _of understanding_ remains. Thus there is some pleasure even in understanding that a range of diverse physical symptoms mark the onset of a terminal illness, or in grasping that one's spouse has been unfaithful, to take to examples where what is understood causes us great sorrow. But still, in either case there is a new understanding, and that in itself brings a pleasurable satisfaction. However, that pleasure is hardly noticeable when it is overshadowed by what we come to understand (cf. Salmon 1989, p. 90).

Understanding as we conceive it pretheoretically is thus at least a discrete, surprising, and pleasant psychological state. In her recent discussion of the phenomenology of scientific explanation, Alison Gopnik (1998) arrives at an account of what "it is like to... have an explanation" (p. 108) that coincides at many points with the account I've presented. What Gopnik calls the "aha" of explanation is, she says, surprising, satisfying, pleasurable, and "strikingly domain general" (pp. 108-110). Gopnik argues that the distinctive and enjoyable phenomenology associated with having an explanation is a byproduct of selection; specifically, the promise of explanation functions to motivate the activation of our "theory-formation system," which in turn contributes to fitness, in the same way that the promise of orgasm motivates intercourse, participation in which contributes to fitness. The details (and some of the main points) of Gopnik's more general account are at odds with what I argue here, but with respect to our accounts of the effect of explanation there is, I think, considerable agreement.

A psychological state very similar to what I have characterized as understanding has been the subject of a series of "insight" studies conducted by Janet Metcalfe and David Wiebe (1986a, 1986b, 1987). In one experiment Metcalfe and Wiebe asked subjects to solve five high school algebra problems, one at a time, and at fifteen-second intervals while they worked indicate "their feeling of warmth (i.e., their perceived closeness of solution)" by placing a slash to the left or the right of a 3 cm scale, depending upon whether they judged themselves to be near or far, respectively, from the solution. The subjects were then presented with a series of five "insight" problems, that is, problems which researchers have historically regarded as requiring insight or illumination for solution (the algebra problems, in contrast, were deemed “grind-out-the-solution” problems demanding no significant insight). In one insight problem (from deBono 1969), for example, subjects were presented with the array



and asked to “show how you can move three circles to get the triangle [formed by the dots] to point to the bottom of the page” (1987, p. 245).

In Metcalfe and Wiebe’s study, algebra and insight problems were associated with different patterns of warmth ratings respectively. Subjects tended to judge themselves far from the solution of an insight problem until seconds before they actually solved it, while subjects judged themselves progressively nearer the solution as they solved an algebra problem. Patterns of warmth ratings for each problem type in the last minute of problem-solving are shown in Figure 1. Here the difference between insight and non-insight problems is graphically represented; for the former but not the latter the solution comes unexpectedly. The trend suggested by Figure 1 is borne out under statistical analysis; non-insight

-------------------------------

Figure 1 about here

-------------------------------

problems were significantly correlated with feeling of warmth patterns that gradually shifted to the right, and insight problems with patterns that shifted abruptly, $p \leq .05$.[vi] Metcalfe and Wiebe conclude that "the phenomenology of insight problem solution [is] characterized by a sudden, unforeseen flash of illumination" (238).

One central contention of the account of explanation offered here is that the "sudden, unforeseen flash of illumination" identified by Metcalfe and Wiebe just is understanding as it was characterized above. Metcalfe and Wiebe's way of describing their data suggests as much. Indeed, Metcalfe and Wiebe motivate this work by associating insight with scientific discovery and, following Robert Sternberg, with "significant and exceptional intellectual accomplishment -- for example, major scientific discoveries, new and important inventions, and new and significant understandings of major literary, philosophical, and similar work" (Sternberg 1985, p. 282, cited in Metcalfe and Wiebe 1987, p. 238). This coincidence of insight as Metcalfe and Wiebe locate it and the aim of scientific explanation as we described it above does not imply that Metcalfe and Wiebe's subjects were engaging in scientific explanations; plainly, they were not. Our understanding of the process that results in insight or understanding will have to be refined if we are to build a theory of explanation on this basis. The fact that insight and understanding are discrete states of very similar phenomenology is only the motivation for that refinement.

The first step in appreciating the theory of understanding suggested by the proposed identification of insight and understanding is a consideration of Metcalfe and Wiebe's study in the broader context of the study of human problem-solving. Following Newell and Simon (1972), problem-solving has typically been understood as a search through a "problem-space." Problem-spaces are defined in terms of an "initial state" $S_I$; a desired "goal state" $S_G$; a set of "operators" $\{O_1, O_2, O_3, \ldots O_N\}$ which applied to one state bring about another; and a set of "path constraints," for example, that the goal state be obtained with the least cost or in the least number of steps (Holyoak 1992, pp. 269-270). A problem-space can be represented with circles for states and arrows for operators. Solving the problem represented in Figure 2, for example, is a matter of two successive applications

-------------------------

Figure 2 about here

-------------------------

of $O_2$ (which brings about State 3) followed by an application of $O_1$ (which brings about $S_G$). Representing a problem in this way makes vivid the idea of problem solving as a search; it also vindicates metaphors we use for problem-solving, metaphors like "hitting a dead end," "going roundabout to the solution," or "going in circles."

The construal of problems as searches in a problem-space suggests a number of problem-solving strategies. One, described by Newell and Simon (1972), directs a solver to calculate the "distance" between the goal state and the present state and apply operators which decrease this distance. Such a "means-ends" analysis may characterize how we solve a large class of problems, but, as Metcalfe and Wiebe point out, it is an inadequate account of the approach subjects take to the problems characterized above as insight problems. If subjects' warmth ratings reflect the distance between a current state and the goal state, then solving an insight problem looks more like a dramatic leap to the solution than a methodical navigation through a problem-space. Indeed, Metcalfe and Wiebe's work on insight problems has led them, and others, to question whether problem solving is always a matter of searching a problem-space. Keith Holyoak, for example, suggests that Metcalfe and Wiebe's experiments show that problem solving is sometimes a matter of defining an ill-defined problem or of restructuring a problem-space (Holyoak 1992, pp. 286-287), and Metcalfe and Wiebe note that their data support the view that insight problems are solved by "some nonanalytical, sudden process" (p. 239). The sudden arrival at a solution to an insight problem is, on these interpretations, a result of abruptly seeing the problem differently such that a solution is almost immediately available (Holyoak 1995, pp. 286-287; Metcalfe and Wiebe pp. 239, 243; Cf. Chi 1992).

This is, however, not the only interpretation of Metcalfe and Wiebe's data, nor is it the most plausible one. Metcalfe's and Wiebe's data allows for the interpretation that subjects merely have no conscious access to the strategies they in fact use to solve insight problems, and this is consistent with taking solutions to insight problems to be navigations of a problem-space rather than giant leaps across, or dramatic restructurings of, that space. Indeed, this interpretation coheres with Metcalfe and Wiebe's (1987, p. 242) conclusion elsewhere that subjects have no privileged access to information that enables them to predict their success at solving insight problems. Moreover, even if we allow that the characteristic feature of insight problems is that solving them involves discovering or restructuring a problem-space, that restructuring itself might be modeled as a search in a broader space, a "meta" problem-space. There is therefore less reason to adopt Metcalfe and Wiebe's interpretation of their experiments. Rather than involving dramatic restructurings or leaps of logic, the characteristic feature of insight problems is merely that their solution comes, for the solver, unexpectedly.

I have argued that insight and understanding ought to be regarded as the same phenomenon, and that the cognitive tasks that elicit insight should therefore be understood as searches through a problem-space, albeit searches distinguished by the solver's lack of conscious access the solution (i.e. $S_G$), however imminent. Conjoined with the identification of understanding and insight, we arrive at the view that that understanding is a discrete psychological state, phenomenologically surprising and pleasant, produced in the wake of solving an insight problem. How then might this rudimentary theory of understanding underwrite a theory of explanation?

Metcalfe and Wiebe's insight problems are not explanation-seeking why-questions, but explanation-seeking why-questions can be cast as insight problems of a certain form. The second central contention of the problem-solving account of explanation is that explanation is the solving of an insight problem which has been formed on the basis of an explanation-seeking why-question. Call such problems "explanation-seeking problems." Combining the representation of problems shown in Figure 2 with the formal erotetic apparatus from Section 1, we can help ourselves to a perspicuous representation of this second contention. Let $\langle P_K, X, R \rangle$ be an explanation-seeking why-question, and let $R = \{\langle A, \langle P_K, X \rangle\rangle\}$, where A is a true proposition. Although this explanation request mentions R, and A may even be a member of the background knowledge K, a request for an explanation is typically made in contexts in which the fact that A bears R to $\langle P_K, X \rangle$ is either not recognized or not appreciated. Giving the explanation is at least a matter of singling $P_K$ out from among the rest of X by means of discovering the proposition which bears R to $P_K$.[vii] The explanation-seeking problem and its solution can be represented as a problem-space, in which the initial state $S_I$ is associated with the expression '<?, <$P_K$, X>>' (where '?' is a variable taking propositions as values) and $S_G$ is associated with the expression '<A, <$P_K$, X>>'. The replacement of the names of various propositions for the variable in '<?, <$P_K$, X>>' comprise the various operators; thus intermediate states are associated with such expressions as '<B, <$P_K$, X>>', '<C, <$P_K$, X>>', etc. The member of R relevant to the question at hand represented in the problem-space as well, as $S_G$. A portion of a problem-space for an explanation-seeking problem is represented in Figure 3. Unlike

----------------------------------

Figure 3 about here.

----------------------------------

the problem-space depicted in Figure 3, typical explanation-seeking problems will have more states and thus more complex problem-spaces. But, as in this simple case, each will pose some specific version of the scientific question '<?, <$P_K$, X>>' as an expression to be "solved," i.e., to be completed by identifying a proposition bearing R to <$P_K$, X>.

To summarize: on the problem-solving account of scientific explanation explanations are navigations of those problem-spaces associated with explanation-seeking problems. Scientific understanding is the payoff for accomplishing an explanation, which is to say, it is the psychological state attained as a result of solving an insight problem. As should be clear, understanding is not on this account an additional sort of knowledge, but neither is it an entirely subjective feature of scientific practice.

In the following section I will consider some objections and consequences of this account, including how it handles two tasks which any theory of scientific explanation must address -- the problem of rejected requests for explanation and the problem of explanatory asymmetry.

## 4. Rejections and Asymmetries

A theory of explanation must account for legitimate *rejections* of putative requests for explanation, and it must handle *asymmetries* of explanation, exemplified usually in the case of the flagpole and its shadow (van Fraassen 1980, pp. 132-134; cf. Kitcher and Salmon 1987, pp. 315-317). How well does the problem-solving account do with these problems?

Consider rejections first. The problem-solving account construes a request for explanation as the setting of a problem, defined in erotetic terms. For each problem then there is a topic $P_K$, a contrast class X, and a relation R. The problem-solving account thus enjoys the considerable success of van Fraassen's account in ruling out inappropriate requests for explanation. Specifically, we can say that an explanation request is *appropriate* in a given context with background knowledge K on the problem-solving account only if $P_K$ is true on K, the other members of X are false on K, and K does not entail that no A bears R to <$P_K$, X>. By analogy with the erotetic account, (a), (b), and (c) can be regarded as presuppositions of an explanation-seeking problem.

For explanation-seeking problems where (a) or (b) are not satisfied but (c) is we have a weakened but wholly familiar sense of explanation, namely the sense in which a falsehood is explained by some other facts. The problem-solving account of explanation accommodates such cases, for when (a) or (b) are not satisfied, but (c) is, there is still a problem to be solved, that is, there is still a well-defined problem-space, the successful navigation of which provides understanding. Accordingly, let us call explanation-seeking problems for which (a) or (b) are not satisfied, but (c) is, fictional explanation-seeking problems, and their solutions fictional explanations. There is no sense of explanation available when (c) is not satisfied, for the requirement that our background knowledge not entail that no A bears R to $<P_K, X>$ translates on the problem-solving account to the requirement that K not entail that the explanation-seeking problem has no solution. If (c) is not satisfied we have an insoluble problem, which in the sense employed here does not qualify as a problem at all.

The imposition of (a)-(c) is necessary, but not sufficient, if the problem-solving account is to handle for legitimate rejections of explanation requests. If the passage from initial to goal state in an explanation-seeking problem is so simple as to be a search in only a technical "Pickwickian" sense, then there is no understanding, and no explanation. Certainly this is true in the degenerate case, in which initial and goal state are identical, but slightly richer problem-spaces with two nodes or one or two operations can, similarly, be expected to provide no understanding upon solution. This suggests adding to (a)-(c)

(d′) The (explanation-seeking) problem is a nontrivial insight problem.

where insight problems are defined in the way suggested in the previous section. The triviality of a problem can be explicated in terms of its problem-space; more trivial problems will presumably have fewer intermediate states and fewer basic operators available to move between those states, and thus will have overall less complex paths to the goal state. The analysis of triviality can be tackled empirically as well, with error rates and time to solution studies as a function of problem-space complexity. Problem-spaces of a complexity associated with very low times to solution and error rates would then be identified as trivial. This empirical research is required if the problem-solving account is to be articulated in detail; prior to it we have only our intuitions to guide us in sorting trivial from non-trivial problems. However, in the cases which will concern us I will take it as given that explanation-seeking problems with only one or two or operators are clearly trivial.

We should also note here that the triviality of problems is presumably partly independent of their character as insight problems; there are, for example, complex non-insight problems. The problem-solving account does conjecture that no trivial problems are insight problems, however. Thus aside from empirical studies of when a problem becomes trivial, the problem-solving account suggests empirical studies of insight, such as the relation of insight to problem complexity just noted and the question of whether all appropriate explanation-seeking problems are also insight problems, as the problem-solving account claims. Although the problem-solving account may be refuted by these studies, I regard its reliance on them as an advantage for a theory of explanation and understanding, both of which are human endeavors.

On the basis of (d′) we can diagnose the cases presented by Kitcher and Salmon, cases which they claim show that R must be an objective relevance relation. Recall that in their examples the lack of constraint placed on R by van Fraassen is exploited to show how to construct what is, on van Fraassen's account, an exemplary explanation of $P_K$ on the basis of A for any true $P_K$ and A in K. Only with the addition of (d) to the conditions an explanation-seeking why-question must meet if it is to arise, they claim, can this reductio be avoided.

On the problem-solving account, as we have just seen, being a maximally telling answer to an explanation-seeking why-question is not enough to be the core of a proper explanation. The answer must also be arrived at as the solution to an insight problem. Otherwise, there is according to the problem-solving account no understanding and no explanation. But in Kitcher and Salmon's examples a non-trivial search for the proposition bearing R to $<P_K, X>$ is precluded by Kitcher and Salmon's explicit presentation of this proposition as the question's answer in the course of describing the example. It is trivial to determine what proposition bears the "astral influence" relation to the proposition that JFK died on 11/22/63 when we are told at the start that R consists of "ordered pairs of descriptions of the positions of the stars and planets at the time of a person's birth and propositions about that person's fate" (Kitcher and Salmon 1987, p. 322). From this initial state it is a small step to the goal state, in which A is seen to stand in R to $<P_K, X>$. In terms of the problem-solving account, in this example Kitcher and Salmon present a trivial explanation-seeking problem, one in which the move from initial to goal state has no intervening states. The nature of R, astral influence in this case, is quite beside the point.

Consider another of Kitcher and Salmon's examples, in which they aim to show that a particular worry for Hempel's "covering law" account of explanation (Hempel 1965, p. 339) resurfaces on van Fraassen's account. Let $P_K$ be that Horace is bald as $P_K$, X be $\{P_K, \sim P_K\}$, A the conjunctive proposition that Horace is a member of the Greenbury School Board and all members of the Greenbury School Board are bald, and R

> the relation of Greenbury-school-board-derivation that holds between A and $P_K$ just in case A is a conjunction of propositions one of whose conjuncts is the proposition that Horace belongs to the Greenbury school Board, $P_K$ is derivable from A, and there is no conjunct in A that could be deleted while still enabling $P_K$ to be derivable from the result. (1987, 327)

Provided A and $P_K$ are true members of K, Horace's membership on the Greenbury School Board figures essentially in an exemplary explanation of his baldness. But from the point of view of the problem-solving approach, the relevant explanation-seeking problem is trivial. As in the previous case, the problem is presented such that the answer's core, A, and the information that A *is* the answer to the question, are given explicitly. The task of completing the expression $<?, <P_K, X>>$ requires a search of one step, from initial state to goal state. This is why "solving" the explanation-seeking problem (or, in erotetic terms, answering the explanation-seeking why-question) produces no understanding.

The problem-solving account agrees with Kitcher and Salmon that these cases contain no explanations, but disagrees with their diagnosis. The difference can be expressed in more general terms. On the problem-solving account, explanation is *dynamic* -- it requires intellectual work, in the form of motion from an initial to a goal state in a non-trivial problem-space. Understanding is the product of such work. In these terms, the Kitcher/Salmon cases fail as explanations not because they employ illegitimate relevance relations, but because in them the dynamic element required for explanation is missing, due to the triviality of the problems.

It is important to note that the diagnosis of Kitcher and Salmon's examples just offered does not turn only on the fact that A is included in the presentation of the explanation-seeking problem. The point is that the problems posed in their examples are trivial; and the inclusion of A in the setting of the explanation-seeking question, while necessary, is not sufficient to make a problem trivial. Thus the problem-solving account avoids objections which turn on explanations with explanans and explananda with which are already quite familiar.

Behind this objection lurks another. The problem-solving account locates explanation in that process which produces understanding, and it identifies this process as the solving of an explanation-seeking insight problem. But, goes the objection, scientific explanation so construed is at odds with scientific explanation as we understand it. Presumably the "sudden, unforeseen flash of illumination" had upon solving an insight problem is less prominent upon solving the same problem subsequent times. But explanations do not decay, no matter how often we apply or cite them. The purported difference is clearest in pedagogical contexts, when a teacher conveys an explanation to students. The teacher traces a problem solution with which she is likely all too familiar. Students following the problem solution for the first time experience the sudden illumination and insight; the teacher does not. Are we to say that the teacher no longer has an explanation, or that what was once an explanation for her has ceased to be one? Clearly not, goes the objection; explanation is thus not the same as solving a problem.

There is some truth to this objection, for surely the experience of solving an explanation-seeking problem is more vibrant and illuminating for students solving it for the first time than for a teacher retracing familiar steps. However, it is not appropriate to describe subsequent solvings of an insight problem as devoid of the sudden illumination that characterizes a first solution, and it may even be the case that subsequent solvings retain a substantial portion of the initial phenomenology. This seems to be the case for sample insight problems used in Metcalfe and Wiebe's study. The psychological "charge" of the sudden illumination one has when seeing which three dots to move in the diagram above, and where to move them, can be recaptured by recalling the problem and its solution. Even the weariest science teacher, I propose, can in a similar fashion recapture the sudden illumination associated with an explanation, and this indeed is what keeps it an explanation.

Conversely, we should recognize that what strikes us at one time as an explanation may in time strike us as something less than an explanation -- for example, a simple description of the world which does not contribute to our understanding of it at all. Initially, for example, we may explain why this penny conducts electricity by noting that it is made of copper, that all copper conducts electricity, and that these facts entail that the penny conducts electricity. In time we may see this same structure as nothing more than an assemblage of facts.

The response I have suggested to this objection works on two fronts: it suggests that explanations need not loose their phenomenological character over repeated use, and on the other hand that over many uses explanations may loose their distinctive explanatory character. Jointly these two replies show us how to reconcile the dynamic character of explanation as the problem-solving account describes it with explanation's static aspects.

At this point we can also note the manner in which the problem-solving account is a pragmatic theory of explanation. Explanation-seeking problems are to be rejected when they fail to satisfy (a)-(d'). The condition (d') is effectively independent of the relevance relation specified in the explanation-seeking problem, and so on the problem-solving account nearly any relevance relation can figure in an appropriate explanation-seeking problem. In particular, whether a relevance relation describes an objective relation between aspects of the world, as Kitcher and Salmon require, is independent of whether it might figure in a non-trivial problem. The problem-solving account is thus a pragmatic account of explanation in the sense outlined at the start of this paper, for on it whether a particular relation of explanatory relevance is appropriate in a given context depends upon contingent facts about the desires, interests, and capacities of the people who produce and consume the explanation. The problem-solving account does not address relevance relations as such, but it is concerned instead with whether they appear as components of non-trivial problems. Moreover, the facts about the consumers and producers of explanations to which it makes reference are presumably fairly deep and universal facts about human problem-solving, as well as more transient facts about current aims, desires, or tastes. Thus it is not the case that on the problem-solving account explanation is "whatever S counts as an explanation," or "whatever satisfies S's curiosity". The problem-solving account rules out no proposed relations of explanatory relevance, in principle, and it does refer to wholly contingent features of human cognition. But it is not an "anything goes" theory of explanation.

Let us turn to the asymmetries of explanation. Sometimes one feature of the world is used to explain another, but not vice-versa. Hempel (1965, p. 352) noted that a pendulum's length explained its period while its period did not explain its length, but the more famous example, attributed to Sylvain

Bromberger (see Salmon 1989, 47), is of a flagpole and its shadow: the height of a flagpole (along with facts about light, the position of the sun, etc.) explains the length of its shadow, but the length of the shadow (along with the same facts) does not explain the flagpole's height. The task facing a theory of explanation is to track this difference; i.e., to reconstruct the explanation in one direction, the intuitively sanctioned one, but not the other.

As simple as the task sounds, it could not be accomplished in Hempel's covering law account of explanation. In that account a simple exchange of premise and conclusion in the relevant argument suffices to turn the derivation of the statement asserting the length of the shadow from statements describing the height of the flagpole, other initial conditions, and various physical laws, into an equally acceptable derivation of the height of the tower from the length of the shadow. The covering law account was unable to recognize only the former explanation as legitimate; thus it offered two symmetrical explanations where it was expected to track a purported asymmetry (see Salmon 1989, p. 47, and 1998, and also Richardson (1994) for accounts of the challenge posed to the covering law account by this and other asymmetries of explanation). How does the problem-solving account fare with the same example?

Explanatory asymmetry arises only when (1) a theory presents us with explanations which are, on that theory, sufficiently (and perhaps even formally) similar to warrant being called symmetrical, and in which (2) these two explanations are *prima facie* of dramatically different explanatory worth. Thus there are two distinct strategies for responding to the problem when it arises. If (2) is granted (in a particular case), then the theory of explanation at hand must be altered so that the explanations the theory presented as symmetrical are no longer presented as such by the theory. For example, the moral many are inclined to drawn from the flagpole case is that the relation of explanatory relevance is causation, and that the covering law account should be modified so that the derivation of the flagpole's height from its shadow (and some laws and initial conditions) no longer counts as an explanation, on the grounds that it does not respect the causal facts of the matter (e.g. Humphreys 1989).[viii]

Attractive as this strategy may be, it is not available to the problem-solving account. To see this, we must first appreciate that, as Kitcher and Salmon show, an "explanation" of the flagpole's height on the basis of the length of its shadow can be easily constructed within van Fraassen's account. Let $P_K$ be a proposition ascribing to the flagpole its actual height, X a set of false propositions ascribing other heights, A a proposition ascribing to the shadow its actual length, and R the relation of "censored Hempelian derivation -- a relation that holds between A and $\langle P_K, X\rangle$ just in case there is a D-N argument that derives $P_K$ from A plus additional premises in [a subset of K] K(Q)" (Kitcher and Salmon 1987, 328; see Richardson (1995) for a critique). Then the proposition ascribing to the shadow its actual length is the core of the maximally telling answer to the question of why the flagpole is as high as it is; i.e., the shadow's length explains the flagpole's height.

The explanation-seeking problem founded on the question $\langle P_K, X, R\rangle$ will not obviously be inappropriate. The conditions (a)-(c) are satisfied, by hypothesis. There is good reason to think (d′) is satisfied as well, not because the explanation-seeking problem involves the use of complex relations and computation (which it does), but because if (d′) were not satisfied, then (d′) would not be satisfied for the explanation-seeking problem which sought the explanation of the shadow's length on the basis of the flagpole's height. As explanation-seeking problems, these two are symmetrical. Of course, (d') could turn out to not be satisfied by either problem, and in various ways they could (each) be rendered trivial. The point is that, on the problem-solving account, the two stand or fall together.

The alternative response to asymmetry cases, and the response I claim is the right one in defense of the problem-solving account, is to accept (1) and deny (2); that is, to claim that the explanations the theory regards as symmetrical do in fact have equivalent explanatory value.[ix] It should be noted that it is easy to overlook the explanatory equivalence of symmetrical explanations if we appraise them, inappropriately, in light of pairs of different, <u>a</u>symmetrical, explanations. So, for example, if we let the appeal of a causal explanation of the length of the flagpole's shadow in terms of its height influence our comparison of two symmetrical covering law explanations, both of which cite inferential rather than causal connections, and if consequently we regard the covering law explanation of the shadow's length on the basis of the flagpole's height as having more explanatory power than its partner, we create the impression of a problem of explanatory asymmetry where no such problem may exist. The two covering law explanations may be of exactly the same explanatory worth. This point is easily lost sight of when we recognize only a single relation of explanatory relevance.

This clears the way for the problem-solving account to handle explanatory asymmetries. If we take care to compare only explanations which are genuinely symmetrical on the problem-solving account, then we can accept what the account suggests, namely that both these explanations <u>are</u> explanations -- neither is to be rejected. If a determination of the length of the shadow given its height (and other assorted facts) explains that length on the problem-solving account, then in the same way and to the same extent the shadow's length explains the flagpole's height. It is quite consistent with this to claim that neither explanation provides much in the way of understanding, and to prefer to either one an explanation which appeals to causal relations. But this is not to show that the problem-solving account fails to accommodate an explanatory asymmetry.

## 5. Conclusion

This presentation and defense of the problem-solving account of scientific explanation has considered several issues surrounding explanation, including issues, like the character of understanding and the relation between understanding and explanation, that have, in my view, been neglected in philosophical examinations of explanation.

Still, if the problem-solving account is to be regarded as a serious candidate theory of scientific explanation, much philosophical work remains. I have not compared the problem-solving account with causal theories or unificationist theories, for example, and there is the range of empirical work on insight and complexity, mentioned in Section 3 that the account requires. In this paper my aim has been to set forth and motivate a new and thoroughly pragmatic theory of scientific explanation and argue that this new account answers to the initial (albeit quite severe) hurdles we set for any such theory. That it survives these severe tests, answers the objections raised previously against pragmatic accounts, and (most of all) does so while forging a deep link between explanation and understanding, is reason enough, I hope, to warrant its further study from different historical, sociological, and philosophical perspectives.

## Endnotes

[i] Like many theorists of scientific explanation, I presume in this paper that the aim of (scientific) explanation is (scientific) understanding. On my view, the converse claim -- that understanding is always the result of explanation -- is denied; understanding, as we will see, can and does arise in non-scientific contexts. For a stronger denial of this converse claim, to wit that we sometimes understand something that we cannot explain (but for which a request for explanation is appropriate), see Rescher (1970, pp. 133-134), cited in Salmon (1989, p. 93).

[ii] Alternatively, the import of (d) could be made a condition upon answers to Q, with the result that questions making reference to a pseudorelevance relation would arise but have no answer. There is some unclarity in Kitcher and Salmon (1987) on this point, for they describe (d) as an additional condition "on answers to why-questions" (1987, p. 322), yet their notation suggests that what they offer is a new presupposition of Q to be satisfied if Q arises in a context. The latter option is in the spirit of their admonition against "silly questions" (p. 322) and is how Salmon presents (d) (1989, p. 143).

[iii] As Kitcher and Salmon (1987, p. 326) note, van Fraassen is not sympathetic to pragmatism so construed either.

[iv] A narrower sense of understanding, usually termed ‘verstehen’, has figured prominently in discussions of the social sciences since Dilthey (see Dilthey (1961) and Winch (1990) for representative discussions, and Salmon, M. (1989) for an overview). Understanding will come in for a more detailed characterization below, but it is worth noting now that understanding in the sense employed here is to be distinguished from verstehen at least by being applicable to natural phenomena beyond human actions and by being characteristic of all sciences, indeed, all inquiry.

[v] Other philosophical discussions of understanding are no more promising in this regard. Hempel’s many references to “scientific understanding” contrasted to “empathetic understanding” (e.g. 1965, pp. 161-163, 239-240, 257-258, p. 329) were followed by suggestive but cryptic explications of the former notion. Salmon’s recent (1998) discussion of understanding offers a taxonomy of types of understanding but does not venture toward an account of what ties these types together, i.e., toward a theory of understanding. Achinstein (1983), on the other hand, proposes that it is a necessary and sufficient condition for A to understand q that there exists a

proposition p which is both a "complete content-giving proposition with respect to Q" and known by A to be a correct answer to Q (1983, p. 42; see also Salmon 1989, pp. 147-148). But, as Salmon (1989, p. 147) points out, the notion of a "correct answer" remains unexplicated. So too, then does "understanding" on Achinstein's account.

For a recent and far more fecund discussion of understanding, motivated by work in developmental psychology, see Gopnik (1998), discussed below.

[vi] Subject's "feeling of knowing," as measured by either their presolution estimation of the probability that they would solve a given problem or their ranking of problems according to which ones they felt they were most likely to solve within four minutes, was predictive of actual performance for non-insight problems, but not for insight problems. "The idea that subjects may have privileged access to idiosyncratic information that makes them especially able to predict their own performance," Metcalfe and Wiebe conclude, "was overwhelmingly wrong in this experiment" (1987, p. 242).

[vii] Further constraints, along the lines that the answer be "telling" (see Section 1), might also be imposed.

[viii] A similar result obtains within van Fraassen's account when R is limited to causation. Then the question asking for a

causal account of the shadow's length presumably has as the core of its answer a proposition ascribing to the flagpole its (correct) height, while the question about the flagpole's height has at the core of its answer not a proposition that ascribes to the shadow its length (for this does not bear the right causal relation to the height of the flagpole), but one which cites various facts involving the causal history of the flagpole's construction. Thus the two questions receive asymmetrical answers.

[ix] As Kitcher and Salmon (1987, p. 316) note, this approach is suggested by Hempel in his response to a variant of the flagpole scenario posed of his covering law account; see Hempel (1965), pp. 352-353.

## References

Achinstein, P. (1983) The Nature of Explanation (Oxford: Oxford University Press).

Chi, M. (1992) 'Conceptual Change Within and Across Ontological Categories: Examples from Learning and Discovery in Science', in R. Giere (ed), Cognitive Models of Science (Minneapolis, Minn.: University of Minnesota Press), pp. 129-186.

Churchland, P. M. (1989) 'On the Nature of Explanation: A PDP Approach', in P. Churchland, A Neurocomputational Perspective (Cambridge: Bradford/MIT Press), pp. 197-230.

deBono, E. (1969) The Mechanism of Mind (New York, NY: Penguin).

Dilthey, W. (1961) Meaning in History H.P. Rickman (ed.) (London: Allen and Unwin).

Friedman, M. (1974) 'Explanation and Understanding', Journal of Philosophy 71, 5-19. Reprinted in J. Pitt (ed.) (1988), Theories of Explanation (Oxford: Oxford University Press), pp. 188-198.

Gopnik, A. (1998) 'Explanation as Orgasm', Mind and Machines 8, 101-118.

Hempel, C. G. (1965) 'Aspects of Scientific Explanation', in C.G. Hempel, Aspects of Scientific Explanation (New York, NY: Free Press), pp. 331-496.

Holyoak, K. (1995) 'Problem Solving', in E. E. Smith and D. N. Osherson (eds.), Thinking: An Invitation to Cognitive Science, Volume 3 (Cambridge, Ma.: Bradford/MIT Press), pp. 267-296.

Humphreys, P. (1989) The Chances of Explanation: Causal Explanation in the Social, Medical, and Physical Sciences (Princeton, N.J.: Princeton University Press).

Kitcher, P. S. and Salmon, W. (1987) 'Van Fraassen on Explanation', Journal of Philosophy 84, 315-330. Reprinted in W. Salmon, Causality and Explanation (Oxford: Oxford University Press), pp. 178-190.

Kitcher, P. S. (1976) 'Explanation, Conjunction, and Unification', Journal of Philosophy 73, 207-212.

Kitcher, P. S. (1981) 'Explanatory Unification', Philosophy of Science 48, 507-531.

Kitcher, P. S. (1985) 'Two Approaches to Explanation', Journal of Philosophy 72, 632-639.

Kitcher, P. S. (1993) Advancement of Science (Oxford: Oxford University Press).

Laudan, L. (1990) 'Normative Naturalism', Philosophy of Science 57, 44-59.

Metcalfe, J. (1986a) 'Feeling of Knowing in Memory and Problem Solving', Journal of Experimental Psychology: Learning, Memory, and Cognition 12, 288-294.

Metcalfe, J. (1986b) 'Premonitions of Insight Predict Impending Error', Journal of Experimental Psychology: Learning, Memory, and Cognition 12, 623-634.

Metcalfe, J., and Wiebe, D. (1987) 'Intuition In Insight and Noninsight Problem Solving', Memory and Cognition 15, 238-246.

Newell, A. and Simon, H. (1972) Human Problem Solving (Englewood Cliffs, N.J.: Prentice Hall).

Rescher, N. (1970) Scientific Explanation (New York: Free Press).

Richardson, A. (1995) 'Explanation: Pragmatics and Asymmetry', Philosophical Studies 80, 109-129.

Salmon, M. (1989) 'Explanation in the Social Sciences', in P. Kitcher and W. Salmon (eds.), Scientific Explanation. (Minneapolis, Minn.: University of Minnesota Press), pp. 384-409.

Salmon, W. (1989) 'Four Decades of Scientific Explanation', in P. S. Kitcher and W. Salmon (eds.), Scientific Explanation. (Minnesota: University of Minnesota Press), pp. 3-219.

Salmon, W. (1998) 'The Importance of Scientific Understanding', in W. Salmon, Causality and Explanation (Oxford: Oxford University Press), pp. 79-91.

Salmon, W. (1998) Causality and Explanation (Oxford: Oxford University Press).

Sternberg, R. J. (1985) Beyond I.Q. (Cambridge, Ma.: Cambridge University Press).

Van Fraassen, B. (1977) 'The Pragmatics of Explanation', American Philosophical Quarterly 14, 143-150.

Van Fraassen, B. (1980) The Scientific Image (Oxford: The Clarendon Press).

Winch, P. (1990) The Very Idea of a Social Science and its Relation to Philosophy, Second Edition (Atlantic Highlands, N.J.: Humanities Press International, Inc.).