Matti Sintonen

Explanation: In Search of the Rationale

The interrogative construal of explanations is one of the earliest analytical devices in the philosophy of science. The general idea is still with us, in part because Hempel and Oppenheim endorsed it in their classic formal account. But although their work starts with the claim that explaining phenomena amounts to answering why-questions, the interrogative perspective did not find its way to either the analysandum or the analysans. And it is fair to say that despite intensive discussion there still is no consensus over the scope and credentials of the perspective. One conceivable reason for this lack of progress is that we haven't had, until recently, a well-developed erotetic logic to deal with questions and answers. Another consideration has kept the doors closed to influences from another direction. Taking the idea of questions and answers literally invites us to look at explanations as anthropomorphic goods, produced and consumed by real-life people. But this brings metatheory to the vicinity of pragmatics, an area which was simply off limits for Hempel and Oppenheim.

The story of pragmatics is a Cinderella story. For quite some time she has had access to the court. However, there still are a number of advisors in the court who keep whispering that the shoe didn't quite fit to begin with. Apart from the objection that pragmatics brings in an undesired element of subjectivity there is the complaint that it has failed to create sufficient order in some notoriously messy corners in the court. I shall argue, in a largely informal fashion, that Cinderella manages her daily routines with grace. The rather loose interrogative idea ramifies into distinct insights, from the logic of questions to the logic and pragmatics of conversation and inquiry at large. Roughly, I shall try to establish that where the former fails to bring philosophical illumination, the latter provides am-

This paper was written while I enjoyed a Fulbright Research Grant in Boston Center for the Philosophy and History of Science. I want to thank the Finland-U.S. Educational Exchange Commission for the financial support and Boston University for its hospitality during my stay. Some passages of this paper are based on my article "On the Logic of Why-questions," printed in *PSA 1984*, Vol. 1, pp. 168–76. (East Lansing, Michigan: The Philosophy of Science Association).

ple consolation. Let me just briefly outline why I think a pragmatic interrogative perspective is needed to restore the lost rationale of scientific explanation.

First, scientific inquiry is goal-directed and rational activity which aims at the acquisition of certain cognitive objectives or goals, such as information, truth, and explanatory power. The interrogative model looks into the conditions under which an answer, intended to promote such objectives, in fact does so. I may fail to convince the philosophical advisors who think that these conditions have little to contribute toward explicating specifically scientific explanation. But it is worth a try, for a number of general reasons. To start with, it is not clear that there are sweeping differences between good everyday explanations and good scientific ones, although there are between silly everyday explanations and good scientific ones. But equally important, airing questions, answers, and explanations is a primordial mode of action, and we should not think that a philosophical account which draws on this insight is any worse off just for that.

Second, to be successful in promoting cognitive goals questions and answers must be to the point. My claim is that the syntactic and semantic mold in which most previous accounts of explanation have been cast creates difficulties which can be overcome, in a quite natural way, by referring to some pragmatic features of questions and answers. For instance, modeling explanations through questions and answers brings in a needed dynamic aspect to explanations. Another such obstacle has to do with the identity of questions: there usually is more to a question than meets the ear, and for an answer to be relevant it must respect this contextual residue. The point to emphasize here is not that syntactic-cum-semantic explicates won't do—that has been clear for a while—but, rather, that to the extent they do they do because they manage to capture pragmatic restrictions by syntactic and semantic means.

Third, purely formal models are incapable of accounting for the variety of explanation-seeking questions, and they remain silent on how questions arise. It might be thought that this opens an opportunity for the logic of questions which insists that the first question always is if the question in fact arises. I shall argue (in the next section) that the logic of questions is too impoverished to be of much assistance, but sketch (in sections 4–6) a pragmatic account of inquiry which is rich enough to give some compensation. Coupling a weak logic with a rich theory of inquiry brings one virtue of Cinderella to full light: questions provide a passage through which salient parts of our more or less amorphous body of background knowledge surface to our consciousness, to constitute objects of wonder and bewilderment.

1. Why-Questions

Let me turn to the first obstacle in the way of a more rigorous interrogative model of explanation. Now that we do have a logic of questions we could perhaps

use it to throw light on the notion of explanation. To explore the proposal we need some terminology and notation. To begin with we distinguish between propositional questions (such as yes-no-questions) which take complete propositions (or sets of propositions) as answers and wh-questions (what-, where-, who-questions) which receive singular terms as answers.

To proceed via examples, let us represent the logical forms of the wh-questions "What is the color of copper?" and "Who authorized the arms deal?" by (1) and (2), respectively,

(1) (?x) (x is the color of copper),

(2) (?x) (x authorized the arms deal),

and call the quantifier expressions in them *interrogative operators* and the open sentences following them the *matrices* of the questions. Now there is a simple relationship between the logical forms of the questions and their presuppositions, i.e., the existential generalizations

(3) (Ex)(x is the color of copper),

(4) (Ex)(x authorized the arms deal),

for questions (1) and (2) have (direct) answers only if (3) and (4) are true. Finally, there is an easy way to tell what a direct answer to a wh-question is: it is simply a substitution instance of the matrix of the question.

Wh-questions have the nice feature that they allow us to use epistemic logic to study conditions under which an answer is conclusive for an inquirer H. An answer is conclusive for H if H is, after hearing it, in a position to say, truly, "I know what (who, where, etc.) x is." A correct direct answer does not automatically fulfill this condition. For instance a (unique) definite description, say "The head of the National Security Council (NSC)," to H's question (2) is not, if H does not know who the head is (or what NSC is). The model provides a tool for the study of answerhood because there is a simple condition under which a direct answer satisfies the questioner. It does so, namely, when and only when the answer together with the inquirer's background knowledge suffices to entail the state of affairs he desires, described by "I know who authorized the arms deal." Clearly, an answer is conclusive only if it does not give rise to such further questions as "But who is the head?" or "But what is NSC?"

There are alternative ways of representing questions, answers and presuppositions, but they all suffice to highlight a difficulty for our project. The idea works where the questioner knows, when putting the question, what counts as an answer. This requirement can be understood in a stronger and a weaker sense. The stronger amounts to an ability to enumerate all possible substitution instances, as in yes-no- (and which-) questions: here the answerer's task boils down to the choice of one of the displayed alternatives. The weaker sense is satisfied when

the questioner knows precisely what *types* of substitution instances would count as direct answers, and is (usually) fulfilled by wh-questions. An answer to a what color-question must specify a color, an answer to a who-question a person, and an answer to a where-question a place to count as an answer of the right type. There may of course be other contextual demands which further narrow down possible substitution instances, but the general observation holds: mere knowledge of language and its categories suffices to guarantee that both the questioner and the answerer know precisely what would count as an answer.[1]

Unfortunately this is not true of why-questions or, more exactly, of the contextually open (if not ill-defined) why-questions which cry for explanatory answers. Granted that explanations are answers to why-questions, there is no logic for why-questions to match in precision the logic of wh-questions. To appreciate the difference, let us try to apply the idea to why-questions, such as the much discussed "Why did the radiator of my car crack?" There is no problem in casting the question into the canonical notation. We can represent the question which arises for H by "(?p)E(p)," where the interrogative operator "(?p)" ranges over causes or reasons and E is the presupposition of the question, i.e., "The radiator of my car cracked." The question, then, has the logical form:

(5) (?p) (The radiator of my car cracked because p).

Direct answers to the question are substitution instances of its matrix "E(p)," and they have the form "The radiator of my car cracked because p," or its truncated form "Because p." Now if I am knowledgeable enough I am able to think of a number of singular facts which, together with some known general laws governing temperatures and volumes, would be sufficient to entail the presupposition of (5). Then the availability of a potential explanation for the event would amount to the fact that I am in a position to enumerate a set of potential causes $p_1, p_2, \ldots, p_n$, any one of which would be a filling of the right type. But this would mean that I would have been able to turn a why-question into a which-question, and the burden of the answer would have been reduced to singling out one alternative.

But although some why-questions have this form, the most interesting and troublesome ones do not. Search for such an explanation is not search for evidence to determine which one of a set of well-defined potential answers is the actual one. In the interesting cases the interrogative operator ("Why?") ranges over causes or, more generally, reasons, the identity of which is contextual in the extreme. Whether (5) in fact is a covert which-question cannot be read from the sentence alone (it depends on the background knowledge of the inquirer). When it came to delineating answers to a scientific question like "Why do the equinoxes precess?," raised by the ancients, the strong knowledgeability condition was not fulfilled. And to the extent the weak one was, it was not because the interrogative operator had a type tag glued to it. It hardly requires pointing out that when we

come to puzzles in quantum mechanics things get out of hand: we encounter questions in which we are completely at a loss to infer, from the interrogative operator alone, what would count as an answer.

I am not suggesting that all why-questions are open-ended, only that some are. And although I shall qualify the account later by suggesting possible category restrictions on why-questions, the value of a logic of questions to any concrete type of inquiry is in proportion to the severity of the constraints on possible answers in that inquiry. And it is noteworthy also that the following anomalies associate with some why-questions. First, it would be tempting to model their presuppositions along those of wh-questions, so that the presupposition of (5) would be

(6) (Ep) (The radiator of my car cracked because p).

i.e., the existential claim that there is a reason or cause for my car radiator's cracking. The reason why erotetic logicians have tilted toward such a rendering of why-questions is that they generally hold that a question has an answer only if its presupposition is true. However, although I am inclined to think that (6) is true (it certainly would be bad news if it wasn't), I am less prone to think that (7) is the presupposition of (8), that is, that there is a direct answer to (8) only if (7) is true:

(7) (Ep)(This uranium atom split because p).

(8) (?p)(This uranium atom split because p).

Now it might be thought that this result merely indicates that (7) does not really arise, and that it only allows corrective answers such as "Because it just did." Be this as it may, it is clear that the choice hinges on knowledge of physics (and maybe metaphysics), not on knowledge of language and the logical form of why-questions.[2]

Second, the following elementary result for wh-questions fails for some why-questions. One attractive view of knowledge has it that someone knows who a person is only if he can tell that person apart from others (in the jargon of this view, if he can identify this individual in the various possible worlds not ruled out by the rest of his knowledge). But I might be puzzling over a why-question and know for sure that I fail to be in an analogous position. If the presupposition of the why-question (that is, the syntactically chosen counterpart of the question-sentence) flatly contradicts what I believe (or believed!), I know that some of my background knowledge is false. It is inconceivable that I could be so baffled over a who-question.

2. A Thin Logic of Questions

It appears, then, that we are at an impasse: maybe the reason why the interrogative model of explanation has had difficulties getting airborne is that there is

no logic of why-questions. Here I want to retract a bit, and ask what happens if we take the pragmatic and intentional idiom seriously. We have an explainee H who addresses a question, and an explainer S, also a member of the scientific community, who intends, by uttering a linguistic expression u, to get the explainee H to understand why E.

For the general idea to work we need to assume that members of the scientific community have a common language in which questions and answers can be phrased, common contextual background beliefs and interests which delineate admissible answers and provide clues for interpretation, as well as common conventions which, if not enable, at least make it easier for speakers and hearers to identify one another's intentions. We could, more specifically, resort to Grice's Cooperation Principle and to the associated more specific maxims, such as the maxims of relevance and truthfulness. All these assumptions can be easily justified, for we have already agreed that scientific inquiry is a cooperative and rational search for such epistemic goods as truth, information content, and explanatory power. So with these pragmatic elements at hand we arrive at a new five-placed analysandum "S explained to H why E by uttering u in a problem context P," instead of Hempel's two-placed "{T, C} is an explanans for E."[3] But to counter the philosophical worries briefly mentioned we must show that these conditions bring in something new. Do they?

The role of u, as intended by S, is to cause in H an epistemic change vis-à-vis the explanandum E. When successful it works along time-honored communicative strategies: S relies on H's knowledge of the meaning of u (if it has conventional meaning) and on whatever means there are for getting H to recognize that u is intended as an answer to H's question. But there are further conditions for the truth of the pragmatic locution which give flesh to the interrogative proposal. First, we saw that there always is more to a question than meets the ear, and for u to bring about the desired epistemic change it must represent a right type of answer to "(?p)E(p)." Second, u must bring in something H did not know already. Third, an answerer S who cooperates commits himself to the principle that u is to the point and does not burden H with information which has nothing to do with the topic. Fourth, the answer expressed by u must be factually relevant: E must be true *because* the answer is. Finally, u must have the intended effect, that is, change H's epistemic state.[4] The five types of condition under which u brings about an improvement in H's epistemic state vis-à-vis E formulate distinct requirements of relevance. In the sequel I shall adopt a notational shortcut and talk, instead of the answer expressed by u, simply of the answer u.

More needs to be said about all of these, but I shall begin with the requirements that u must be news to H, and that the news must pertain to the question at hand, i.e., to why E. Let us start with the commonplace that explanation has to do with the improvement of knowledge. Some H may perceive E as problematic because it, say, contradicts (or is improbable relative to) H's background knowledge. This

means that there is, first, a (more or less) harmonious *initial epistemic state K* in which H is peacefully unaware of forthcoming epistemic trouble. Second, there is the intermediate *problem state* K_E which characteristically results when the observation that E throws the initial state into disequilibrium. Believing is a cognitive and not conative attitude, so that H is forced to admit that E and forced to admit that he doesn't know why E. Finally there is the happy *end state* K_u in which harmony is restored (at least locally) by bringing in the answer u: H then knows both that E and why E. The three epistemic states reveal in fact a dialogue pattern: the problem state gives rise to H's es-question "(?p)E(p)" and the answer is intended to cause in H the end state K_u.

Now to some specifics. It has proved to be dishearteningly difficult to design a satisfactory formal model of deductive explanation within the static confines of Hempel and Oppenheim's original proposal (hence called H-O). My suggestion is that one reason for this difficulty is that the formal models purport to capture some pragmatic, dynamic and contextual features of questions and answers *by purely syntactic and semantic means*.

Here are some examples. According to H-O a pair {T, C} is an explanans for a (singular) E if and only if (1), T is a theory, (2), C is a true singular sentence, (3), {T, C} $\vdash$E, and (4), there is a class K of basic sentences such that K is compatible with T, K $\vdash$C, but not K $\vdash$E.[5] Here condition (4) is a syntactic restriction which rules out arguments in which the singular premise C entails the explanandum E. Complete self-explanations in which C and E are identical are good examples. But what is wrong with them? They are answers in which an inquirer H is offered a piece of information which he already knows.

Consider next what Hempel and Oppenheim called self-evidencing explanations, such as

(9) T: (x)Mx
C: Ma $\supset$ Pa

E: Pa

According to Hempel and Oppenheim (9) is unacceptable because, given that T is true, the only way of verifying the singular premise C goes via the verification of the explanandum E. But why the requirement of independent verification? Look at the matter from a dynamic point of view. In the problem state H knows that E but not why. His puzzlement needs for alleviation an answer which gives grounds to believe (expect) that E, but not all answers which, together with H's prior knowledge, entail E will do. No answer which is derivable from what H already knows in K_E can improve his epistemic state with respect to E. The requirement of independent verification points to an important fact: for the explanans to be conclusive for H, it must be such that, when added to H's initial state, it would give independent grounds to expect that E.

One intuition behind H-O was that all scientific explanations involve laws, hence the requirement that E be derivable from {T, C} but not from C alone. But what is the rationale behind this? Let us take another look at (9). Robert Ackerman has used it to demonstrate what he calls the *trivialization principle*: an explanation is unacceptable if the explanans can be used to explain anything.[6] And clearly if (9) is all right, then we could replace Pa by anything (excepting Ma and $-$Ma) and get an explanation. Now there is an obvious interrogative intuition behind shunning such versatility, for the law in (9) completely fails to address the specific question 'Why E?' Laws, then, are needed to secure what was called above the factual relevance condition, i.e., that E is true *because* the answer is.

Another example is one of the famous counterexamples to H-O produced by Eberle, Kaplan, and Montague.[7] Assuming that we have a true fundamental law (x)Fx and a true singular sentence E, we can derive from (x)Fx another fundamental law T which "explains" E, even if the original law and E have no predicates in common:

(10) T: (x)(y)(Fx v (Gy $\supset$ Hy))
C: (Fb v $\sim$Ga) $\supset$ Ha

E: Ha

Now (10) satisfies the conditions of H-O but is nevertheless unacceptable as explanation. One way to block it was proposed by Jaegwon Kim: add to H-O the requirement that no conjunct in the conjunctive normal form of C be entailed by E, and (10) becomes unacceptable. Hempel welcomed the device but added: it would be desirable to have a non-ad hoc justification for it, one which is not based on the brute fact that the requirement blocks the derivation but finds support in "the rationale of scientific explanation."[8]

The rationale is to improve H's knowledge vis-à-vis the explanandum. Just recall the basic dynamic feature of the question-answer model. Three epistemic states are involved. In the problem state H already knows that E, and therefore knows everything E entails. In the example at hand E entails all of the conjuncts of the conjunctive normal form of C, so that C cannot bring news to H. More specifically, it cannot bring anything relevant about E, viz., independent grounds to expect that it is true. What Kim's syntactic requirement does is bring in the intermediate problem state into the assessment of an argument.

That there is a simple interrogative intuition behind the syntactic requirement can also be seen from rival proposals. D. A. Thorpe has given one in terms of MECs. A MEC (Minimum Evidence Class) for a singular sentence S is the smallest class of basic sentences which suffices to verify S. All we need to require, then, is that no T-consistent MEC for C can verify E. So what is the rationale? The MEC-requirement is in accordance with the scientific method, writes Thorpe, because "we would want the theory to mediate or else we would have

no D-N explanation at all."[9] But of course a theory mediates between C and E when C neither entails nor is entailed by E. The interrogative intuition gives the rationale for both: if C entails E, the law is idle and the inquirer H has no reason to believe that E is true *because* C is. Where E entails C the answer C or any part thereof fails to enlighten H in any way at all, because it is already known in K_E.

The requirement that T should mediate between C and E points to one way in which an answer should be relevant. Another one derives from the answerer S's commitment not to flood H with information which does not bear on E (and hence, from the implicature that E is true because the answer is). Consider Henry Kyburg's counterexample to H-O: The sentence E ("This piece of salt dissolved in water") can be "explained" by referring to the general law T ("All hexed samples of salt dissolve in water") and to the singular sentences C_1 ("This sample of salt was placed in water") and C_2 ("This sample is hexed").[10] But even assuming that the sample indeed was hexed (so that the explanans becomes true), $\{C_1, C_2\}$ and C_2 both fail as answers to the question "Why E?" Although both are true, neither is a true answer to the question: it would be false to claim that E is true because the answer is. Clearly both answers flout the general conversational requirement of pertaining to the point.

An analogous requirement rules out redundant information in the explanandum.[11] But these examples should suffice to show that the insight in H-O – that explanations are answers to why-questions – is not buried in any interesting logic of why-questions, but rather in the admittedly thin conversational logic of question-answer sequences: to the extent the syntactic constraints on arguments work they work because they capture unabashedly pragmatic features of discourse.

3. The Epistemic Conception of Explanation

I have all along taken it for granted that the capacity of an answer u to enlighten H resides in its capacity to make the event described by E nomically expected. It is noteworthy that Hempel and Oppenheim baked this rationale into their reading of (singular) why-questions from the outset. The assumption defines what Salmon has called epistemic and Stegmüller informational concept of explanation (in two books dedicated to Hempel).[12] There are a number of considerations which show that the assumption generates trouble, but no consensus over the remedy. As a modest step toward such remedy I shall generate some more trouble, this time from a dynamic epistemic model which steers around problems of relevance. The hope is that this point of view gives us easier access to the root difficulties.

The model is Peter Gardenfors's knowledge situations model which has it that the purely logical tools employed in H-O (and most other formal models) are not rich enough for an adequate explication of explanation.[13] Instead, Gardenfors in-

sists that singular explanations, deductive and statistical alike, are evaluated with respect to the knowledge situations (or epistemic states) of a rational agent, and, more generally, that determining whether something is an explanation depends on the pragmatic context. Now a knowledge situation is a quintuple of the form

$$(11)\ K = \langle U, W, \{I_w\}_{w\in W}, \{P_w\}_{w\in W}, B\rangle.$$

where U is a domain of individuals, W a set of possible worlds (compatible with what H knows), $\{I_w\}_{w\in W}$ an interpretation function which assigns each individual constant of the language L an element from U and each predicate of L a set of individuals. $\{P_w\}_{w\in W}$ is a probability measure for the subsets of individuals in W. It assigns each singular sentence of L its value in every world $w\in W$. B is H's *belief function* which assigns for each subset V of W H's degree of belief that the actual world is an element of V. Now H is characteristically ignorant of the objective probabilities of different individuals' having various properties. To obtain the probability with which H expects an individual a to have the property G Gardenfors resorts to mixed probabilities: the values of the measure P_w are multiplied by H's subjective probability that w is the actual world. As a result we get the following formula for the expected probability of an individual a to be G (for a finite V):

$$(12)\ P_V(G) = \frac{1}{B(V)_{w\in V}} P_w(I_w(G))\ B(\{w\}).$$

Assume then that some singular event described by the sentence E is unexpected for H. This means that H's initial epistemic state assigns the sentence E which describes the event a low belief value. Assume now that H nevertheless observes that E. This observation causes (brings about) a new belief state K_E and gives rise to the why-question "(?p)E(p)." What kind of answer would be conclusive for H? Gardenfors does not require that the answer raises the belief value to the vicinity of 1, nor above .5, but he does insist that it raises it to some degree at any rate. But raises relative to what? H already knows in K_E that E is true, therefore its value is 1 and cannot be raised. Gardenfors opts for another candidate, the initial epistemic state. So an explanans is conclusive for H if and only if it raises the value of E relative to its value in the initial epistemic state?

The proposal gives a precise measure for the capacity of an answer u (or explanans {T, C}) to make the truth of a singular explanandum E expected. The degree of rational tension created by E is its belief value in the problem state minus its belief value in the initial epistemic state. The relief offered by a given explanans—its local explanatory power with respect to E—is measured as follows: first add, by making necessary subtractions and alterations, the answer to the initial epistemic state. Then determine the expected value of E in the end state and subtract from this value the value of E in the initial one. The result tells you how much relief you got.

The model also recommends a strategy for finding the best explanans: it is the one which, when added to the initial knowledge state, maximizes the belief value of E. It also bars partial and complete self-explanations, for adding to the initial epistemic state general (and singular) sentences which are already included in the intermediate one do not improve H's epistemic state vis-à-vis E. Making the relevance of an answer u dependent on the inquirer's knowledge states K and K_E also explains why the acceptability of the answer "Because he was a heavy smoker" to the question "Why did John get lung cancer?" varies from person to person. It is acceptable only for an explainee who does not already know that John was a heavy smoker (or that smoking causes lung cancer). Knowledge situations and the associated belief functions introduce a fine-grained criterion of identity for why-questions: there simply is no one-one mapping from the (surface) why-questions to the explananda, and an inquirer who knows that John was a heavy smoker expects to hear an answer which tells him why, among heavy smokers, just John got the disease.[14]

But the model is incomplete at best. First, assuming that making things expected, or at least raising belief value, is all there is to explanation, why should H in the intermediate epistemic state bother searching for an explanation, *for the belief value of the explanandum E in that state already is the maximum possible?* Put in slightly different terms, if the belief value of the explanandum E in K_E is one, and if increase in belief value earmarks growth of knowledge and understanding, what could possibly count as improvement? Moreover, in statistical explanation at least the belief value of E is *higher* in the problem state K_E than it is in the allegedly harmonious end state, for the simple reason that statistical explanations do not raise the belief value of the explanandum E to one. Therefore, if high or at least higher belief value is the goal, you should stick to the problem state in which you know that E but not why.

Second, there is an oddity in the epistemic account which raises related questions. In discussing – and dismissing – the view that explanations must reduce or link the puzzling event to something with which the questioner is already familiar, and that the familiar needs no explaining, Hempel cites a number of familiar facts, such as tides, lightnings, and even the fact that it is dark at night, which have been subjected to explanation.[15] The examples are puzzling because the familiar facts and the regularity are highly expected and have high belief values. They do not fit the epistemic model for the simple reason that whatever explanations are offered the belief values of the sentences "There are tides," "There are thunders" and "It is dark at night" are not inched up a bit. Again, improvement in our knowledge cannot be more expectedness.

One way out of this difficulty would be to deny that the explananda in the problem situations do have the maximum value one. The reason for this move is easy to appreciate. E is singled out as an explanandum because the inquirer H's background knowledge develops to the point when, suddenly, the fact that E becomes

problematic. From a purely logical point of view one could think that this merely shows that E does not, after all, have belief value one. But this way out – Bradley's way because he blamed facts when they conflicted with his high principles – is not open to an H who wishes to stay sensitive to experience. And it *is* implausible to think that the degree of belief of the proposition "It is dark at night" is decreased just because there are theoretical considerations which show that it should be low. H is far more confident of the truth of this common observation than of the background theory which made it problematic.

Another possible way out seizes on H's epistemic split-mindedness and appeals to H's sense of epistemic integrity. What has happened is that H has been saddled with two belief functions which deliver different verdicts on a salient E, and a respectable way out is to move to a third state in which the discrepancy is removed. This way out is, I agree, the right one. The trouble is that it finds little support in the official epistemic doctrine which views explanations as arguments within an interpreted language: it is a static view, geared to maximizing belief value, and it does not tell why an inquirer H should keep an eye on past states.

The difficulties gain in intensity when we move away from H-O to the explanation of regularities and laws, to an area which was not the concern of Gardenfors's model.[16] There is a sense in which regularities might be unexpected – say, when they conflict with known laws. And one could imagine that surprise elimination could cover some theoretical explanations. But whether or not this would fit the intent of the model, there are why-questions which are more resistant to the treatment. First, there are well-entrenched brute regularities which are not surprising or unexpected in any sense but which still pop up as question marks. Second, it often happens in science that we are able to find similarities and formal analogies between phenomena or types of phenomena, such as the inverse square form of laws found in different but perhaps partly overlapping physical theories. Now there may be no uncertainty about the descriptive adequacy of the empirical laws involved – they have high belief values. But of course such formal analogies cry for explanations, in terms of common physical mechanisms, if possible. Whether or not such mechanisms can be found is of course a matter of empirical inquiry. In any case it is clear that we have here a why-question 'Why do the empirical laws have a similar form?' which does not arise from the low belief value or unexpectedness of the generalizations but from the entirely different desire to bring the two established regularities under one roof.

Nor is this the end of troubles for the epistemic model. It is not reasonable for an agent to worry about all isolated beliefs or sentences which happen to have low belief values, or over all changes of belief values toward the worse. It is essential to the mental health of an agent that he be able to live with such tensions, on pain of being paralyzed. For an E to become a genuine problem it is also necessary that E matters, i.e., that it is cognitively central enough. One more shortcoming in the epistemic model is that it provides no account at all why some events in

the ocean of unexpected and unrelated events need explanation and others do not.[17]

4. Theory Nets and Explanatory Commitments

A need for an explanation arises when there is a particular why- (or how-, or sometimes what-) question but no satisfactory answer. We can also agree that a belief state K_E in which you intuitively speaking know that E but don't know why is a state in which something is found to be missing. But to make these commonplaces more than platitudes we need an account which is able to provide the inquirer H motivation for worrying not just over low belief values but also stranded beliefs and, more generally, over forms of disintegrity weaker than inconsistency. The knowledge states model of Gardenfors contains some of the right ingredients, for there are the three epistemic states K, K_E, and K_u. However, what we need is an additional note on epistemic integrity: how do these knowledge states hang together? What makes an inquirer in K_E responsible for his commitments in $K-$and what are these commitments, anyway? The lead we shall follow is this: the crucial impetus must reside in the tension created by the claim of a theory (or more generally, a body of knowledge) to handle an area of experience, and its (temporary or chronic) incapability of substantiating this claim.

This is a handsome order, and we should not expect that anyone can deliver all ingredients at once. But in setting out to fill it we do not need to start from scratch, for there is a theory notion which is designed to meet some of these desiderata and which can be augmented to meet the rest. I have in mind the so-called structuralist theory-notion. I shall give an informal account of it, and confine my discussion to those features of the notion that are of importance to our immediate concerns.[18] A theory, or more precisely a *theory-element* T is an ordered pair $<K,I>$ in which the *core* K is a conceptual apparatus, and I a set of (classes of) *intended applications*. The core K is a quadruple $<M_{pp}, M_p, M, C>$ in which M_{pp} stands for the potential partial models of the theory, roughly, non-theoretically described objects or constellations of objects, and M for the laws of the theory. M_p represents those objects which have enough structure of the right sort so that it makes sense to ask whether they can be enriched with theoretical functions so as to satisfy the laws M of the theory. Finally, C is a set of constraints whose role in the core is to guarantee that functions which appear in distinct intended applications receive same values.

The range of intended applications I (a subset of M_{pp}) is in turned delineated intentionally through some paradigmatic exemplars. This means that there is no strict set of necessary and sufficient conditions for an object to fall into the domain of the theory. Rather, there are measures of similarity which require that any object or constellation of objects sufficiently similar to one of the exemplars is in the domain of responsibility of the theory. A theory-core can now be used to make

empirical claims, viz., claims that some intended applications are models of T, that is, belong to M. The claim of the theory-element at large is the claim that the entire range of its intended applications are structures which can be supplied with theoretical functions in a way which makes them models of T.

Now theory-elements in science often form larger "maxi-theories" or *theory-nets* N, finite sequences of theory-elements connected with one another by various types of intertheory relations. A theory-net has a fundamental or basic theory-element $<K_o, I_o>$ and a number of specialized theory-elements $<K_i, I_i> \epsilon N$, introduced to make more specific claims about some more limited classes of applications ($I_i \subset I_o$). The fundamental core of the theory-net may then give rise to several branches of specializations, and the result may be a hierarchial tree-structure. To give an example, the fundamental core of classical particle mechanics comprised the three Newtonian laws, but these were immediately specialized to distance-depending forces and then, by adding the gravitation law, into a specialized theory-element for gravitation phenomena. Newton's theory was a branching structure in that the distance-depending theory-element was also specialized to nongravitational phenomena (such as oscillating systems), and there were still other branches.[19]

Reference to the specialized theory-elements discovered by a certain time already contains a pragmatic allusion, and it can be made more explicit still by relativizing theory-elements and nets to scientific communities and time-intervals: then we could pick out a set of nontheoretical structures, a subset of I, to which a scientific community SC intends to apply the core at a time (or during a period). Clearly, such intentions to tackle explanatory and other questions reflect not just pocketed achievements but also beliefs and hopes for future net-expansions. One way to codify a scientific community SC's aspirations is to think that it possesses not just a conceptual cutlery K and food for thought I but also other items which, though perhaps more difficult to access formally, are of almost equal importance. These include the paradigmatic exemplars I_p which generate the set of intended applications, as well as the analogies, explanatory ideals, and cognitive values of the scientific community. We can assume that the latter are bequeathed to members of SC through a paradigmatic theory net N_p which shows how the core K_o has been applied thus far, and where it ought to be expanded.

We have now ended up with what has been called a *Kuhn-theory* $<K_o, I_p, N_p, \mathcal{M}>$, where K_o is the fundamental core, I_p its set of paradigmatic exemplars, and N_p an associated paradigmatic theory-net. Now one trouble with any such pragmatized theory-notion is that there are, at any given time, several alternative (and incompatible) directions in which an extant theory-net could be expanded. Some constraints are therefore needed. I shall follow a previous suggestion by the structuralists and represent these contextual beliefs and hopes of SC by a restriction function which weeds out (from among the logically possible ones) expansions

which are contrary to the spirit of the theory, and which gestures toward promising ones.

The pragmatized notion of a Kuhn-theory allows us to study the development of a theory in time. When a Kuhn-theory is proposed, only some of the intended applications, namely those in I_p, have been shown to be models of the theory. It is a task of later generations to refine and expand the theory-net to cover the remaining envisioned but so far unexamined or unsuccessfully examined applications. A *theory-evolution* represents such historical development: it is a finite sequence of $< K_o, I_p >$-based theory-nets $N^1, N^2, \ldots$ such that each N^{i+1} contains at least one theory-element obtained by specialization from an element in the historically preceding theory-net N^i.

It is no part of my argument that the set-theoretic setting in which the structuralist theory notion was cast provides a philosophically impeccable foundation for either the statics or dynamics of theory formation.[20] But representing a theory as a quadruple of a conceptual apparatus, a set of paradigmatic applications, a paradigmatic theory-net and a set of values, analogies and standards embodied in the net has a number of attractions. To begin with, it removes a difficulty with purely formal views which fail to address the problem of how questions arise. If we think that an inquirer H draws on a pragmatically enriched Kuhn-theory, anchored in a set of intended applications, it becomes obvious that H's questions (explananda) do not fall out of the sky: there is a "disciplinary matrix" which gives its holders title to assume that some questions, viz., those which arise from salient nontheoretical structures, are both sound (the presuppositions are true) and motivated while others are not.

Equally important, a Kuhn-theory displays the explanatory commitments of a scientific community, and therefore explains why its member, inquirer H in K_E, is responsible for his commitments in K. H's primary commitment lies in a Kuhn-theory, and since they have a more than ephemeral span they provide a sense of continuity over the specific claims which come and go – therefore H is not at liberty stop in K_E. Note specifically that such a Kuhn-theory could be instrumental in specifying items needed in the knowledge situations model, viz., the domain of individuals U, set of worlds W considered by H to be seriously possible, and belief function of H in a knowledge situation of the type (11). Clearly analogies and heuristic models furnished by a theory-net reflect on H's subjective degree of belief that the actual world is among a given subset of the possible ones.[21]

The pragmatic account also enables us to overcome other problems in the epistemic model. A Kuhn-theory, or, more precisely, a scientific community SC which subscribes to it, characteristically makes a very strong claim, viz., that all of the intended applications picked out by I_p are amenable to treatment by the conceptual cutlery K_o. However, at any given time only some of the intended applications have been confirmed, i.e., only part of the claim has been substantiated. This means that there is a constant tension between the claim of its holders and

the goods delivered thus far. But it also allows us to recover the missing motivation for an attempt to find an answer for an explanandum sentence E, whether singular or general, which already has the maximum belief value 1 in the *initial* epistemic state. For any thus far unexamined or recalcitrant intended application is a challenge of the required form: holders of the theory have known all along that E but haven't known why. Since they are committed to a theory-net which specifically addresses the question by claiming that all intended applications are models, an answer u which fills the lacuna becomes a pressing desideratum.[22]

The proposal also does away with a related anomaly. The epistemic model does not find room for why-questions which arise as queries concerning unaccounted for similarities between two types of phenomena, or as queries over formal analogies between laws and regularities. It is easy to see how such questions arise within a Kuhn-theory, for there are existing special laws which provide analogies (and models) for future net-expansions. A good example of the grip of a formal analogy is provided by the inverse-square form of Newton's law of gravitation, for once it was established it gave rise to well-defined questions in other intended applications, such as electrostatistics. The architectonic beauty of the structuralist theory-notion is now easy to appreciate: already available theory-elements provide heuristic guidance for net-expansions which, if successfully carried out, would have explanatory appeal because the results would fit an already existing pattern. But finding such a law does not always end the inquiry, for there may be a further query about the physical foundations of this common pattern.[23]

Let me next show how this account fits our overall picture. We started with a difficulty with why-questions, viz., with the unavailability of a "logic" which could impose restrictions on the substitution instances of the matrices. And it *is* a living concern for any interrogative account of explanation or inquiry at large that questions have enough structure to rule out some logical possibilities, for otherwise the questioner or inquirer literally does not know what he is looking for. We can now see that there is no reason to despair, for the impoverished logic of why-questions has a rich supporter behind the scene. Although the matrix of a why-question, as it was defined, does not contain strict requirements for admissible fillings (as they do in wh-questions), a disciplinary matrix in the form of a Kuhn-theory gives ample encouragement and reward.

We have now joined the pragmatic idea of questions and answers with a pragmatic account of how questions surface from our more or less amorphous background knowledge. But to the extent a Kuhn-theory provides lacunae of various sizes and shapes we can see that this background knowledge is not all that amorphous, in the mature sciences at least. Normal scientific puzzle-solving presupposes well-defined questions—and that is precisely what a Kuhn-theory gives. Moreover, there is a qualitative change in the types of questions, for the rich theory notion chops the unmanageable why-questions into more manageable wh-

questions and even yes-no questions. A query which starts as a loose why-question concerning an area of experience turns (in the "paradigmatic" phase) into a what question: *what* special laws are needed to govern this particular application? It is a further task to find yes-no- and which-questions that portray specific alternatives and still narrow down admissible substitution instances. There is of course no logic for generating questions—but there is a contextual and nontrivial heuristics.

Notice, next, that the pragmatized theory-notion outlines an account of cognitive centrality which goes some way toward alleviating two related troubles. We noted that not all beliefs with low belief values motivate a search for explanantia. On the other hand only a portion of our background knowledge is relevant for assessing answers, once a salient question is fixed. Van Fraassen has raised the latter problem: assuming that we do have a specific question (with a built-in relevance relation and a contrast class which single out requirements for an answer), there still is the problem of determining the portion of background knowledge which is taken to be relevant in the assessment of goodness. "The evaluation," he writes, "uses only that part of the background information which constitues the general theory about these phenomena, plus other 'auxiliary' facts which are known but which do not imply the fact to be explained." But, he continues, neither he nor others have had much to say about how the portion is delineated. And he concludes that this must be a further contextual factor. Now of course a Kuhn-theory not only singles out questions but also provides a first delineation of background information relevant for answers: clearly all laws and singular facts which belong to the branch of the theory-tree which gave rise to the question are relevant. Similarly, special laws from neighboring branches count as relevant, at least to the extent that they furnish analogies.[24]

5. Pruning the Web of Belief

We now have an interrogative account which makes an inquirer in K_E responsible for his commitments in K. But there is another surprise waiting for us. The *rationale* of an answer u (or a scientific explanans {T, C}), to the extent there is one such thing, is not merely or even mainly to raise the belief value of E, but to synchronize the inquirer's belief states K_E and K. This result is in perfect harmony with the view that explanation is not a local matter in which an explanans conveys some privileged epistemic status, such as familiarity, naturalness, or whatever, to an explanandum. Rather, it is a global matter of reducing the total number of (types of) phenomena that must be accepted as ultimate.[25]

In this section I shall draw attention to one further attraction of Cinderella: embedding the interrogative proposal into the pragmatized theory-notion throws light on explanatory unification. But before seeing how, note two features an analysans should have. First, it should cherish the distinction between singular

facts and kinds of facts. Any genuinely lawlike sentence generalizes over an infinite domain. But as examples gleaned from the history of science show, the sheer number of facts explained is immaterial if not outright ill-defined – what counts is the variety and the number of independently established laws.[26]

Second, the thin logic required that for an answer to be conclusive for H it must contain enough information, but not more, to entail, together with H's background knowledge, the desideratum of H's questions "I know why E." This means that when H's background knowledge expands the answers become shorter. Moreover, if the answerer faces an array of questions, a good strategy, when possible and financially feasible, is to invest in theory. It just might turn out that a number of answers draw on the same concepts and laws. Conversely, it adds to the appeal of a theory if it yields not just a great number of answers, but does so with the minimum of conceptual machinery.

The next step is to show that the proposal comes some way toward securing these features. To begin with, an inquirer H who subscribes to a Kuhn-theory literally inherits a range of intended applications. But the members of this set are not particular structures but classes of applications (formally I_p is a subset of the power set of M_{pp}) and they fall into clearly distinguishable types. There is a clear sense in which Newton's theory brought about unification, for it showed that despite wide variety all instances in intended applications (gravitation phenomena, oscillations, etc.) satisfied the three fundamental laws. Or rather, since we need to distinguish between confirmed and hoped for claims, for each class of applications either there are special laws derived from the fundamental core or there is a premium for deriving them. One ingredient in the degree of unification of a theory-net (at a certain time or during a period) can therefore be assessed with the help of the number of distinct types of intended applications derived from a basic theory-element.[27]

The second desideratum follows from another built-in feature of the structured theory-notion, because, in it, an inquirer equipped with a Kuhn-theory can only avail himself of a limited number of concepts in an attempt to derive special laws (or in an attempt to obtain answers to other questions). The structuralists have stressed that the fundamental core K_0 of a theory-net is a *tool* which is used in the derivation of special laws. Thus a Newtonian working on electromagnetic phenomena finds only some types of special laws admissible and initially plausible. A perspicuous way to put this is to say that an existing theory-net defines a number of asymmetric intertheory relations within its theory-elements, so that each theory-element *presupposes* a number of more fundamental theory-elements.

A recent generalization of intertheory-relations, the notion of a link, goes further still, by extending the prospect of unification beyond theory-nets.[28] On that construal the global structure of science exhibits systematic conceptual connections not just within the elements of a theory-net but also more generally between

theories which, however liberally understood, are not offsprings of the same basic theory-core. Just to give the idea, the concept of a force appears both in classical particle mechanics and Lorentz's electrodynamics, and it must be assumed that at least where the two theories have joint domains the values of the force function are identical.[29] Where an intertheoretic link is asymmetric it can with justification be called a presupposition link. Whether or not the two theory-elements have a common core or conceptual structure it may turn out that the very identification of the set of non-theoretical structures picked out as intended applications of a theory-element T (say, thermodynamics) may require the use of another theory-element T′ (fluid dynamics), but not vice versa.

To see how intertheoretic links between theory-elements help us to understand unification, consider Philip Kitcher's proposal.[30] Crucial for it is the concept of an explanatory store of arguments relative to a set of accepted sentences. Explanations, according to him, are arguments, and scientific theories bring about unification by providing, not unrelated individual arguments, but arguments which exhibit common patterns, imposed by the extra-logical terms available for explanatory purposes in individual arguments. While logic is only interested in the logical form of arguments, theories impose further restrictions. The general problem, then, is to specify the set of patterns which gives the most parsimonious way of generating accepted truths as conclusions.

Furthermore, Kitcher writes, apart from argument patterns there are also what he calls core patterns: some of the arguments which have accepted sentences as conclusions are based on the conclusions of other arguments. Thus, e.g., in Newtonian dynamics there is a core pattern which is used to compute equations of motion. This core pattern is supplemented by patterns of arguments (problem-reducing patterns) used for deriving further conclusions. It may then turn out that all patterns contain a unique core pattern. Clearly the degree of unification of a theory with a single core pattern is greater than the unifying power of a rival with several core patterns. And this must be reflected in the criteria for the best-unifying theory.

Now one consequence of the existence of presupposition links is that a question phrased in a theory T may presuppose prior answers by T', when T' is needed to establish some existential presuppositions of questions in T (say, classical particle mechanics requires for its application prior answers to wh-questions in kinematics). Furthermore, presupposition links have a methodological role in that answers provided by T derive argument patterns from T'. An example, not quite as rigorous as those in physical sciences, can be found in evolutionary theory. Suppose you have the question why island populations exhibit traits different from but clearly similar to those found on the mainland. The evolutionary explanations characteristically refer to the relative geographic isolation of the islands and to the new selective forces operative in them. But to get a satisfactory answer to the question one needs to supplement these empirical observations with a

genetic account of gene distributions, for only such an account can give an answer to the crucial questions of why the advantageous phenotypes survived, and why the selective forces lead to a new species. This shows that answers to questions in systematics presuppose answers by population genetics. Furthermore, population genetics is *the* fundmental core of evolutionary theory, because precisely the same pattern of arguments occurs in paleontology, morphology, and other applications. Whether or not evolutionary studies form a theory-net analogous to those in physical sciences, the concepts and laws of population genetics are needed in all of them.[31]

The account still needs an addendum and a qualification. Both have to do with what it is for an answer to be better. Theory-nets provide a static report on the state of the art, but to explain explanatory progress in evolutionary theory we must resort to theory-evolution as it was defined above. Theory-evolutions can be progressive in three ways at least, and all exhibit increase in unification. First, the claim of a theory-net **N** is the claim that for all $< K_i, I_i > \epsilon N$, I_i can be turned into structures which satisfy the special laws of K_i. One line of progressive theory evolution, then, consists of the substantiation of such claims (and of the redemption of these hopes) by producing new special laws to fill gaps.

Second, there can be refinement in theory-element cores, that is, replacing of existing special laws by more stringent ones, while leaving the sets of intended applications and potential models untouched, as in the various types of theory reduction. Such replacements result in more stringent answers to particular questions, but there is an even more dramatic road to increased unification: finding a new, more restrictive theory-core close to the fundamental element would increase the coherence of the entire edifice, for it would create a new core pattern and hence have an effect on all answers presupposed by it. Third, there is no need to require that the set of intended applications remain fixed during an evolution. One more type of increase in unification results when new classes of applications are added, i.e., a theory-net N_{i+1} replaces a net N_i such that and that $I_i \subset I_{i+1}$, where the new classes were not previously thought to fall within the applications of the theory at all.[32]

The qualification we need has to do with a distinction between providing an answer and providing a good or conclusive—or better or more conclusive—answer. It would be too harsh to say that there can be no answers or even good answers to questions in, say, systematics without prior answers to questions in population genetics. But the more modest comparative claim does hold. A unified theory in which the former questions enjoy support from the latter provides better (more conclusive) answers to all questions, in that it does not leave (or leaves fewer) unanswered questions. It then appears that two outwardly unrelated answers within distinct applications can warm up each other—if they use a pattern derived from a common core.

This introduces a global or holistic element to explanation which needs some

comments. First, to the extent theory-evolution redeems hopes in expansion the fundamental theory-core gains in confirmation. But this gain of course manifests in overall explanatory success, for clearly the fundamental theory-element can be used to provide answers as to why the derived special laws hold. There need not be one type of relationship for all theory-nets for an explanatory relationship to hold. The intertheory-relation which ties two (or more) elements T' and $T(\epsilon \text{ N})$ together can range from reduction (in which, say, the intended applications I' of a macrotheory are correlated with those of a microtheory T, and where whatever T' claims of I' is entailed by what T claims of I) to weaker interpretative relations. And the notion of a link is more versatile still, because it can tie together theories with distinct conceptual structures. But however varied in form, links enhance global coherence, and consequently add to the appeal of the particular answers which emanate from the global structure.[33]

Furthermore, we have received some light on the distinction between theoretical and singular explanation. Although scientific questions often arise over particulars, the concern is not with this or that particular, say the whereabouts of Mercury. Rather, the trouble with the anomalous conduct of Mercury was that it was bad news for theory – it radiated epistemic pangs inward in the web of belief (i.e., upward in the theory-net). Similarly, exciting as the particular discoveries of the geographic distribution of finch populations in the Galapagos Islands were, the particular place, the time, or the species mattered little, except as a stage for good theoretical news. The conclusion to emphasize here is that the answers to the singular queries would have carried little force unless there were reasons to expect that the same pattern was applicable elsewhere – in answering questions about other geographical distributions and other kinds of phenomena.

6. Beyond the Third Dogma of Empiricism

The argument thus far has been that the syntactic-cum-semantic explicates divert our attention from the rationale of explanation. Two findings were instrumental in its recovery, viz., the thin logic of questions (or conversation) and the more bulky structured theory-notion, but both added pragmatic ingredients to the stew. During this search we discovered that unification in its various forms is a fundamental explanatory virtue, and that the rich theory-notion makes it possible for us to explicate that virtue. A number of intriguing questions, both systematic and historical, remain. First, why did it take so long to realize that explanation is a pragmatic affair, and why do we still feel uncomfortable about it? Second, the thin logic of questions, with the annotated constraints on relevance, and the requirement of increase in unification are, in a sense to be made explicit shortly, purely formal. But assuming that they operate in all contexts, are there any pragmatically varying further requirements?

Hempel and Oppenheim's model of scientific explanation grew out of the

logical-empiricist research program which aimed at purging all armchair philosophizing from metascience (and science). This was to be achieved by rational reconstructions of notions such as scientific theory, empirical support, and explanation. The program had as its epistemological hard core the emphasis on experience, but equally constitutive were the values of objectivity and intersubjective testability, designed to rule out not just metaphysics but also all manner of psychologism and anthropomorphism. Logical positivism was not the first philosophy to attempt to set philosophy on the secure path of science, but it was the first to emphasize the systematic use of the conceptual tools of modern logic. If there is anything that captured the spirit of the new scientific way of thought in action, it was the idea that all explications made exclusive use of syntactic and semantic concepts. It is therefore appropriate to follow Stegmüller and label this idea the third dogma of empiricism.[34]

It is obvious from Hempel's early works that these concerns were behind his covering-law model. He wrote, in 1942, that there are two notions of understanding, psychological and theoretical.[35] The psychological (and anthropomorphic) sense is exhibited by a "feeling of empathic familiarity," and it is often conveyed by persuasive metaphors and models. Scientific understanding is brought about by knowledge of facts and regularities which make an event expected, or a regularity a matter of course. One way to capture this theoretical mode of understanding, and to capture it within the required level of conceptual hygiene, was to stick with the dogma and bar reference to persons, scientific communities, and other contextual features of inquiry and language use.

It is remarkable that when the three intended applications, the notions of theory, empirical support, and explanation, ran into trouble there were fingers which pointed to the third dogma. But now that we have gone pragmatic in public, have we thereby admitted, as Michael Scriven has, that the notion of explanation derives its meaning entirely from the subjective and psychologist notion of understanding disdained by the positivists?[36] The answer is an emphatic "No!," for we can, and should, deny the equation of the pragmatic with the psychological and subjective. Although the analysandum "S explained to H why E by uttering u in a problem context P" makes reference to an inquirer's successive epistemic states, the acceptability of an answer does not hinge on any idiosyncratic or anthropomorphic sense of intelligibility. The same is true of Gardenfors's refinement of the epistemic construal: although it incorporates subjective belief functions, no subjectivism or psychologism is involved, for *changes* in belief functions are subject to rigorous normative constraints.

How about the pragmatic elements which enter into the interrogative idea *via* its immersion in an explicitly pragmatized theory-notion? Giving and appraising explanations now becomes an explicitly social affair, because the explananda are circumscribed by a disciplinary matrix, subscribed to by a scientific community at given time. But this relativizing need not lead to relativism. Although the selec-

tion of questions may be up to the scientific community, picking out the answer need not be, in any damaging sense. There may be contextual requirements of type-appropriacy, along with shared analogies, values, and methodological standards, which make the evaluation of answers contextual and paradigm-dependent, but these restrictions on answers need not be taken at face value. There is always the possibility of denying that a question has a direct answer within the bounds of the matrix. It does not, of course, follow from this that there are unique choices to be made.

But there is an element in the enriched interrogative model which enables a more radical departure from the positivist program. The thin logic of questions and answers, as well as the demand for unification, are analogous to principles of rational decision-making in that they lay down constraints on successive epistemic states but remain purposefully silent on the content of individual states. This leaves open the possibility that there are distinct types of inquiry with distinct demands on modes of understanding and standards of intelligibility to supplement the universal but pale procedural canons. This becomes evident once we leave the area of natural science.

To appreciate this possibility we need to go beyond science, narrowly understood. As an introduction, let us take a brief look at the rise and fall of explanatory monism, starting with Aristotle, for he was not just the father of the interrogative model but also a pluralist with respect to questions.[37] Although he started with a description of the "conditions of natural change" he extended the theory of *aitiai* to mathematics, music, biology, and reasons for human action. Aristotle's theory of why-questions, or his theory of explanation, must be seen against his general theory of substance, and the implicit teleology of ancient Greek thought. Details aside, Aristotle thought that each substance had its natural inclinations, and that why-questions contained four clearly discernible, built-in type requirements which codify these inclinations and (types of) causes of aberration. By the advent of modern natural science during the early Renaissance there had been enough changes in the general *Weltanschauung* to make the implicit teleology lose all remaining appeal. As a result other than efficient causes became suspect and not even worthy of the title.

The new notion of explanation championed by David Hume and John Stuart Mill required that all explanations refer to regularities and antecedent events which together allow us to predict events. To cut a long story very short, Hempel and Oppenheim then canonized this view, and applied the new standards of metatheoretic success, the third dogma. Now the dogma not only requires that analyses be conducted within formal languages, it also requires that the formal languages be interpreted in an empiricist language common to all inquiries. The reason there was no need to refer to persons and scientific communities was simply that all theories and explanations were phrased in a language accessible to all, and once you had a complete description of a state of affairs in that language you

had described all there was to an explanandum. But this means that there is essentially but one way to slice the world, and no partisan angles to explananda. Hence the idea that there was a one-one mapping between why-questions and explananda.

The knowledge situations-model of Gardenfors goes beyond this limitation because it individuates explananda, in part at least, in terms of fine-grained belief functions, and explanantia in terms of changes of belief functions. But there are some reasons to think that there is more to an explanandum than is built into knowledge situations in the sense of (11)—there is the possibility that why-questions come in distinct types and that there is no one-one mapping between problem situations and why-questions, either. This would amount to a partial withdrawal of the claim made earlier, viz., that there are no type requirements in why-questions.

If the historical detour above is on the right track, the best candidates for examples are to be found in areas in which the reductive attempt to eliminate all but efficient causes met with greatest resistance, in historical, teleological, and functional explanation, and especially in the explanation of action. I shall take action theory as an example, although parallel stories could be told for other types of explanation.[38] Suppose that you have the question "Why did S go to the market place after lunch?" which, given your knowledge of S's commitments to be in an important meeting at the same time calls for some explaining. A causalist would look for an explanation of the end result of the action, S's going to the market place after lunch, in terms of S's bodily behavior, and beliefs and wants which caused the behavior. S's desire to have fresh fish for dinner and his belief that unless he went to the market place right after lunch he could not have fresh fish for dinner, would do nicely. No two mental cause theories agree on all details, but all agree that if the explanation was correct, it was correct because the want and the belief caused the behavior in S's performing the action, in the ordinary (Humean) sense.

An intentionalist would agree with the causalist that the end result, the explanandum, is explained by wants and beliefs. However, he would disagree over the role of efficient causation in the explanation: on his view actions are understood by construing for them an intelligible background of wants, beliefs, and other determinants such as duties.[39] Two deep disagreements are crucial here. First, by the intentionalist's lights an adequate background story establishes a conceptual, not causal, connection between the explanandum action and its determinants, which means that there is no way to verify their independent existence. Second, the intentionalist insists that he is in the business of understanding intentional action and not bodily behavior. He readily admits that there is the distinct why-question (or perhaps how-question) about the emergence of S's bodily behavior, but when it comes to explaining action we are not interested in it.

The interest of the dispute is not in the outcome but in the initial task of singling

out the explanandum. If the intentionalist is right, the causalist makes a category mistake when he identifies the explanandum. There is, in fact, no one-one mapping between the (surface) why-question "Why did S go to the market place after lunch" and explananda, because the explanandum may be either about S's action or the bodily behavior. In the former case the interrogative why-operator ranges over wants, beliefs, and like determinants as "final causes," in the latter they are construed as efficient ones. This gives us one reason to speak of the revival of Aristotle, for to the extent his and our conceptual schemes mesh, the intentionalist proposal aims at rehabilitating one of his four causes or types of explanatory factor.

There is of course no magic in the number four, nor in a romantic return to Aristotle. The strategy I have favored is to impose a minimum of initial structure on explanations and leave the door open for more specific contextual restrictions. This minimalism is not a counsel of despair, although it is motivated by the fate of explanatory models which have made a commitment to historically changing standards of intelligibility or modes of explanation. And the strategy finds further support in the rules for enlightened debates in which disputants agree to withdraw to a common ground.

This explains why I prefer the phrase factual relevance over causal relevance: the function of "because" in explanatory answers is to tie the explanandum with the explanans, and to indicate that the answer is intended to synchronize H's epistemic states. But this "because" may carry a different force in different types of inquiry. It would greatly unify our total world view if it carried the same force in action theory and science, but it would be unwise to judge the issue in the theory of explanation. And there are further types of becauses, in theoretical linguistics and, say, in questions involving legal liability.[40] However, in all these types of inquiry the thin conversational logic says what an anwer is, and the criterion of unification says what a good answer is. And there is a further feature: in action explanation, too, the explanation is no good if the explanandum is not true because the explanans, or answer, is true. Explanations, then, form a family of more specific notions which arise from the common ground of weak procedural rules and the requirement of factual relevance, i.e., that the explanandum E is true *because* the answer is. There is both variety and similarity provided by genetic kinship between members of the family: they all derive from the parent explicandum "S explained to H why E by uttering u in a problem situation P"

Notes

1. For the logic of questions and the conditions of answerhood, see Hintikka (1976), chapter 3. My indebtedness to Hintikka goes deeper than is outwardly visible in the paper, although the interrogative perspective here developed is not his. I also wish to thank Sylvain Bromberger for linguistic and other evidence (in his [1987] Publication and in personal communication) for the peculiarity of those why-questions which seem to require an explanation for an answer. Roughly, why-questions of that type do not have midsentence traces, either in surface structure or in deep structure. Now the

transparent nature of wh-questions, manifest in the mutual deducibility of the questions and what Bromberger calls their attributive presuppositions (presuppositions to the effect that the substitution instances of the matrix of the question are of a certain type), hinges on the existence of such traces. It follows that a Rational Ignoramus who knows that E but does not know why is unable to figure out, from the logical form of the question alone, what the attributive presupposition is, and, hence, what substitution instance counts as answer. This does not mean, of course, that a Rational Ignoramus does not know what counts as an answer, but it does mean that this knowledge is not of a linguistic nature.

2. See Bromberger (1987) for further discussion.

3. There are rival possible pragmatic analysanda, such as Peter Achinstein's (1983) "S explains q by uttering u," Stegmüller's (1983) three-placed "S explains a to H," and Tuomela's (1980) six-placed "S scientifically explains q to H by producing a linguistic token u in situation C, given P" where P stands for a paradigm or a constellation of group commitments. For these and some other alternatives, see Sintonen (1984a).

4. In the jargon of speech-act theory this means that "explaining" is a perlocutionary act. Achinstein (1983) recognizes both the illocutionary and the perlocutionary senses, but adopts the former. On his view the intended effect of understanding why E, which might or might not follow an illocutionary act, is not part of the act but something brought about by the act. There is no denying that there is a sense of "explain" which does not require that the explainee comes to understand why E. I have wished to adopt the perlocutionary sense because it nicely derives both senses from the same analysandum. Thus the achievement reading of "S explained to H why E by uttering u in a problem context P" might be false even when H correctly identifies the intended answer u, u is not known to H, u is of the right type, and u is to the point, if H cannot accept it (for instance because it is contrary to H's other well-entrenched beliefs). Notice that the locution could be false even if u does have the intended effect—if u in fact does not provide a correct anser. See Achinstein (1983), chapter 2, section 1.

5. Hempel and Oppenheim (1948), pp. 277–78.

6. Ackermann (1965), p. 162–63.

7. Eberle, Kaplan, and Montague (1961).

8. See Kim (1963) and Hempel (1965), p. 295.

9. Thorpe has another justification for the MEC-condition: should it happen that, when T is tested, E does not take place, we would rather blame the relatively uncertain T than the initial conditions. I do not see any general justification for this, however. See Thorpe (1974), p. 191. Gardenfors (1976), p. 430, has formulated a condition on explanations which also secures that T is used in explanation: If the pair (T, C) is an explanation of E, then (a) there are no predicates in E which don't occur in T and (b) there are no predicates in C which don't occur in T. Gardenfors leaves the status of this condition (his C5′) open, however.

10. Kyburg (1965). Kitcher (1981), p. 523, has another explanation for the unacceptability of the argument, namely, that accepting it would lead to an undesirable and avoidable proliferation of argument patterns.

11. For discussion, see Omer (1970), Thorpe (1974), Gardenfors (1976), and Sintonen (1984a).

12. Salmon (1984) and Stegmüller (1983). Hempel and Oppenheim (1948, 246) wrote that the question "Why does the phenomenon occur?" is construed as meaning "according to what general laws and what antecedent conditions does the phenomenon occur?" And it is clear from his writings elsewhere that the function of the law is to make things expected. Salmon appears to take the logical form of the covering-law models, deduction or induction, as constitutive of the epistemic notion. The interrogative proposal generalizes this point: a model of explanation in which answers are designed to more expectedness or higher belief value count as epistemic. I also think it is slightly misleading to make, as Salmon does, erotetic or interrogative models a subset of the epistemic ones, because the rival models, the modal and ontic models, can be given a natural erotetic explication. Indeed it does not seem to me that van Fraassen's (1980) is a purely epistemic one.

13. Gardenfors (1980). In presenting knowledge situations I have however, followed Stegmüller's (1983) exposition. I have briefly discussed the model in Sintonen (1984b), and wish to thank the anonymous referee of the paper for pointing out some difficulties in the epistemic construal.

14. Gardenfors attributes this fine-grained principle of individuation to Hansson (1974). However, Gardenfors only regards why-questions as linguistic indicators of the deeper knowledge situa-

tions. As will become evident in section 6 below, my account is closer to Hansson's original proposal in that it leaves room for type requirements which are irreducible to knowledge as it is spelled out in knowledge situations of the form (11). Van Fraassen (1980) makes this reading of why-questions explicit by baking the pragmatic contrast classes and relevance relations into the questions. However, on his view the pragmatic aspects are determined by openly anthropomorphic goals and interests, and this may well trivialize the problem. For further discussion, see section 6 below.

15. Hempel (1965), pp. 430–31.

16. It is arguable that explaining the familiar fact that it is dark at night already goes beyond explaining singular facts, for we can hardly ask "Why was it dark last night?" There is another worry about Hempel's examples: it is not always clear what the explananda are. Hempel has insisted that the relata in the explanation relation are sentences, and the explanandum phenomena characteristically events. But when we explain tides and lightnings, it is unclear if the explananda are sentences which record that there are tides, or particular occurrences of tides and lightnings, or sentences which describe regularities. I shall not, however, dwell on this.

17. This is only partly true of Gardenfors's model, for he gives an elaborate account of the information relevant for a question, when the question is a demand to make the explanandum less surprising. However, the model does not extend to explain regularities and unrelated events.

18. See Balzer and Sneed (1977) and (1978), Sneed (1976), Stegmüller (1979), and, for more recent developments, Balzer, Moulines, and Sneed (1986).

19. For a structuralist exposition of the Newtonian theory and its developemnt, see Moulines (1979).

20. Initially, the structuralists stressed their view as an alternative to the statement view of theories in which theories are construed as sets of sentences axiomatized in a formal language. It is arguable (and has been argued) that this introduces a false contrast, for there can (and must) be a way to translate set-theoretic talk to more traditional model-theoretic talk by assigning the elements in a theory-core a suitable language. Stegmüller makes a concession to this direction by observing that the difference between the two frameworks is one between a very cumbersome formal framework and an intuitively appealing and simple informal way. I do not wish to take a strong metatheoretical stand here – but it does appear to me that any adequate account should be rich enough to represent not only the structure of theories but also their dynamics and the various kinds of intertheory-relations needed in it. For discussion, see Niiniluoto (1980), Pearce and Rantala (1983), Sintonen (1984a), and Stegmüller (1979).

21. Wesley Salmon thinks that one of the most important uses of analogies is to provide prior probabilities needed, e.g., in an objective Bayesian approach. The account given here is but an attempt to show how analogies and models fit in a pragmatically enriched theory-notion. It is a further task to spell out the precise nature of the inductive logic or other elements required, and how they are used in, say, theory-choice. See Salmon (1984), p. 234.

22. The problem has earlier been treated in connection with the problem-solving model of inquiry. Thus for instance Hattiangadi (1979) has argued that all scientific problems have the form of an inconsistency, whereas Laudan (1977), Lugg (1979), Leplin (1980), and Nickles (1981) have tried to find room for a wider notion of a scientific problem. Both Laudan and Nickles have maintained that a mere logical compatibility of a phenomenon with a theory counts as a problem if there is a "premium" for solving it (Laudan) or if the phenomenon falls in the domain of the theory. The account sketched here translates problem-solving talk to the interrogative idiom and provides, I hope, a fuller account of what a "premium" to solving a problem amounts to. See also Sintonen (1985) for further discussion.

23. Why-questions (and how-questions) thus seem to possess both gosh value (the answers provide intellectual pleasure) and golly value (they generate new accessible questions) in Bromberger's sense. See Bromberger (1985), p. 312.

24. Van Fraassen (1980), p. 147. Gardenfors's (1980, p. 412) knowledge-situations model in fact gives an elegant fine-grained account of the knowledge relevant for a (singular) sentence "Qa," for its belief value is determined by the intersection R of all classes R_i such that "$R_i a$" is true. The belief value B(Qa) is simply P(Q/R) where P is the expected probability function in the knowledge situation. The virture of the more coarse-grained account in terms of the global structure of science is that it accords well with the intuition that the delineation of background knowledge is clearly structured and hierarchically ordered.

25. Friedman (1974) and Hempel (1965), p. 345, where Hempel says that theoretical explanation deepens our understanding by showing that regularities exhibited by a variety of phenomena are "manifestations of a few basic laws."

26. Thagard (1978) shows that it is far from clear that Darwin's theory beat its competitors in deriving more facts, but it was superior in its ability to use a few principles to cover a number of classes of facts, in systematics, embryology, paleontology, morphology, and so forth. Thagard also makes explicit reference to the structuralist idea of intended applications. See also Friedman (1974), pp. 15–16.

27. This gives a rough measure of what I have elsewhere called, following the terminology and intent of William Whewell, the degree of consilience of a theory-net. Unification, it appears, also covers other facets, such as the simplicity of the laws within a particular class of applications. It would go beyond the scope of this paper to go into details, but here we have one reason for the difficulties in attempts to elucidate unification. As I have argued, briefly, in Sintonen (1986), the notion of simplicity covers several distinct intuitions which may be balanced differently in the assessment of the degree of unification or simplicity of a theory.

28. See Moulines (1984) and Balzer, Moulines, and Sneed (1986) for further details.

29. I owe this example and much of the following discussion to Moulines (1984). See also Balzer, Moulines, and Sneed (1986).

30. Kitcher (1981).

31. I owe this pattern of argument and the example to Michael Ruse (1973), chapter 4. Ruse states without reservation that "*population genetics is presupposed by all other evolutionary studies*," and claims that "evolutionary theory is a *unified* theory with population genetics as its presupposed core." Ruse argues that it has a hypothetico-deductive pattern, but it appears to me that the notion of a theory-net which can accommodate a number of distinct kinds of intertheory relationships would be more appropriate.

32. The three types of increase in unification correspond to Stegmüller's three types of evolutionary progress, i.e., confirmational, theoretical, and empirical. See Stegmüller (1979), p. 33.

33. Friedman (1983, section VI.3.) gives a lucid and detailed account of theoretical unification which results when a (characteristically observable) substructure is embedded in a more abstract larger structure. He also notes that unification results in better confirmation: a phenomenological regularity can be well confirmed, but it gains in degree of confirmation if it is derived from a higher-level theory. According to him derivability as such is relatively unimportant; what counts is the interplay between what he calls phenomenological representations and (realistically understood) reduction in time: thus some theoretical assumptions designed to explain (relatively) observable phenomena can be later confirmed by their use in the accommodation of other observable phenomena. This picture is easily translated to queston-theoretic parlance, the way I have sketched. Note, however, that the notions of theory-net and intertheory link allow a more fine-grained classification of the various aspects of unification. For a tentative structuralist classification of intertheory-relations, see Krüger (1980). For unification and confirmation, see also Glymour (1980).

34. See Stegmüller (1983). That logical analysis provided a model to be followed is clearly visible in the Vienna Circle manifesto, *The Scientific Conception of the World: The Vienna Circle*. It quotes (p. 8) Russell's *Our Knowledge of the External world*: the method of logical analysis "has gradually crept into philosophy through the critical scrutiny of mathematics. . . . It represents, I believe, the same kind of advance as was introduced into physics by Galileo: the substitution of piecemeal, detailed and verifiable results for large untested generalities recommended by a certain appeal to imagination."

35. "The Function of General Laws in History." Reprinted in Hempel (1965, pp. 239–40.

36. See Scriven (1975), p.4.

37. According to Julius Moravcsik the popular Aristotelean doctrine of four causes really is "a theory about the structure of explanations." Instead of talking about *aitiai* as causes we should really talk about explanatory factors in a wider sense. See Moravcsik (1974).

38. For a review of the various types of action theory, see Tuomela (1982). Note that there is also a third main type of action theory, the agency theory, which draws on an irreducible type of agent causality.

39. The intentionalist I have in mind is G. H. von Wright, who gives Aristotle credit for the view

that the *aitiai* of human actions are type-distinct from the *aitiai* of things in *rerum natura*. Von Wright has argued that Aristotle had two syllogisms, the theoretical and the practical, with the difference that the theoretical syllogism has as its conclusion a proposition while the practical one has as its conclusion an action (or maybe a proposition which describes an agent's attempt to perform and action). He has also held the very strong view that practical syllogism is a *sui generis* mode of explanation; roughly, it is for human sciences what the covering law is for natural sciences. He has, however, modified his views since. See von Wright (1971).

40. Sylvain Bromberger advances a similar view, with examples from linguistics, in his (1985), p. 321. And it is noteworthy that linguistic arguments against a particular description or explanation standardly refer to its incapability of accounting for this or that generalization. Note also that "because" may carry a different force in different compartments of natural science, as is indicated by the troubles of efficient causation, say, in quantum theory.

References

Achinstein, Peter. 1983. *The Nature of Explanation*. New York and Oxford: The University Press.

Ackermann, R. 1965. Deductive Scientific Explanation, *Philosophy of Science* 32: 155-67.

Balzer, Wolfgang and Sneed, Joseph. 1977 and 1978. Generalized Net Structures of Empirical Theories, *Studia Logica* 36 and 37: 195-211 and 167-194.

——, Ulises Moulines, Carlos, and Sneed, Joseph D. 1986. In The Structure of Empirical Science. *Logic, Methodology and Philosophy of Science VII*, eds. Barcan Marcus et al.. New York: Elsevier, 291-306.

Bromberger, Sylvain. 1965. An Approach to Explanation. In *Analytical Philosophy, 2nd Series*, ed. R. J. Butler. Oxford: Basil Blackwell, 72-103.

——. 1985. On Pragmatic and Scientific Explanation: Comments on Achinstein's and Salmon's Papers. In *PSA 1984*, Volume Two, eds. P. D. Asquith and P. Kitcher. East Lansing, Michigan: The Philosophy of Science Association.

——. 1987. What We Don't Know When We Don't Know Why. In *Scientific Inquiry in Philosophical Perspective*, ed. Nicholas Rescher. Lanham, MD: University Press of America 75-104.

Eberle, R., Kaplan, D., and Montague. R. 1961. Hempel and Oppenheim on Explanation, *Philosophy of Science* 28: 418-28.

Friedman, Michael. 1974. Explanation and Scientific Understanding. *Journal of Philosophy* LVVI, No. 1, 5-19.

——. 1983. *Foundations of Space-Time Theories. Relativistic Physics and Philosophy of Science*. Princeton, NJ: Princeton University Press.

Gardenfors, Peter. 1976. Relevance and Redundancy in Deductive Explanation. *Philosophy of Science* 43: 420-31.

——. 1980. A Pragmatic Approach to Explanation. *Philosophy of Science* 47: 405-23.

Glymour, Clark. 1980, *Theory and Evidence*. Princeton, NJ: Princeton University Press.

Hansson, Bengt. 1974. Explanations—Of What? Stanford University, (Mimeographed).

Hattiangadi, J. N. 1978. The Structure of Problems II, *Philosophy of the Social Sciences* 9: 49-71.

Hempel, Carl. 1965. *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. New York: The Free Press.

——, and Oppenheim, Paul. 1948. Studies in the Logic of Explanation. Reprinted in Hempel (1965), 245-95.

Hintikka, Jaakko. 1976. *The Semantics of Questions and the Questions of Semantics*. North Holland: Acta Philosophica Fennica, Vol. 28, No 4.

Kaplan, David. 1961. Explanation Revisited, *Philosophy of Science* 28: 429-36.

Kim, Jaegwon. 1963. On the Logical Conditions of Deductive Explanation, *Philosophy of Science* 30: 286-91.

Kitcher, Philip. 1981. Explanatory Unification. *Philosophy of Science 48*: 507-31.

Krüger, Lorentz. 1980. Intertheoretic Relations as a Tool for the Rational Reconstruction of Scientific Development, *Studies in History and Philosophy of Science* 11, No. 2, 89-101.

Kyburg, Henry E., Jr. 1965. Discussion: Salmon's Paper. *Philosophy of Science*. 32, no. 2, 147-51.

Laudan, Larry. 1977. *Progress and Its Problems. Towards a Theory of Scientific Growth.* London and Henley: Routledge & Kegan Paul.

Leplin, Jarret. 1980. The Role of Models in Theory Construction. In *Scientific Discovery, Logic, and Rationality*, ed. T. Nickles. Dordrecht, Boston, London: D. Reidel, 267–83.

Lugg, Andrew. 1979. Laudan and the Problem-Solving Approach to Scientific Progress and Rationality. *Philosophy of the Social Sciences 9*: 466–74.

Moravcsik, Julius M.E. 1974. Aristotle on Adequate Explanations. *Synthese* 28: 3–17.

Moulines, Carlos Ulises. 1979. Theory-Nets and the Evolution of Theories: The Example of Newtonian Mechanics, *Synthese* 41: 417–39.

——. 1984. Links, Loops, and the Global Structure of Science, *Philosophia Naturalis. Archiv für Naturphilosophie und die philosophischen Grentzgebiete der exakten Wissenschaften und Wissenschaftsgeschichte*. Band 21, Heft 2–4. Meisenheim/Glan: Verlag Anton Hain.

Nickles, Thomas. 1981. What is a Problem that We May Solve It? *Synthese* 47: 85–118.

Niiniluoto, Ilkka. 1980. The Growth of Theories: Comments on the Structuralist Approach. *Proceedings of the Second International Congress for History and Philosophy of Science, Pisa, 1978.* Dordrecht-Holland: D. Reidel. 3–47.

Omer, I. A. 1970. On the D-N Model of Scientific Explanation, *Philosophy of Science* 37: 417–33.

Pearce, David and Rantala, Veikko. 1983. 'New Foundations for Metatheory'. Synthese 56. Pages 1–26.

Ruse, Michael. 1983. *The Philosophy of Biology*. Atlantic Highlands: Humanities Press.

Salmon, Wesley. 1984. *Scientific Explanation and the Causal Structure of the World.* Princeton: Princeton University Press.

The Scientific Conception of the World: The Vienna Circle. In *Otto Neurath: Empiricism and Sociology*. M. Neurath and R. S. Cohen, eds. Dordrecht, Holland: D. Reidel.

Scriven, Michael. 1975. Causation as Explanation, *Noûs* 9, 3–15.

Sintonen, Matti. 1984a. *The Pragmatics of Scientific Explanation*, Acta Philosophica Fennica, Vol. 37. Helsinki: Societas Philosophica Fennica.

——. 1984b. On the Logic of Why-Questions. In *PSA 1984*, Vol. One, eds. P. D. Asquith and P. Kitcher. East Lansing, Michigan: The Philosophy of Science Association, 168–76.

——. 1985. Separating Problems from their Backgrounds: A Question-Theoretic Proposal. *Communication and Cognition, 18, No 1/2*, 25–49.

——. 1986. Selectivity and Theory Choice. *PSA 1986*, Vol. One, eds. A. Fine and P. Machamer. East Lansing, Michigan: The Philosophy of Science Association, 364–73.

Sneed, Joseph. 1976. Philosophical Problems in the Empirical Science of Science: A Formal Approach. *Erkenntnis* 10: 115–46.

Stegmüller, Wolfgang. 1979. *The Structuralist View of Theories. A Possible Analogue of the Bourbaki Programme in Physical Science.* Berlin, Heidelberg and New York: Springer-Verlag.

——. 1983. *Erklärung, Begründung, Kausalität*. Zweite, verbesserte und erweiterte Auflage. Berlin, Heidelberg, New York: Springer-Verlag.

Thagard, Paul. 1978. The Best Explanation: Criteria for Theory Choice, *The Journal of Philosophy* LXXV: 76–92.

Thorpe, D. A. 1974. The Quartercenternary Model of D-N Explanation, *Philosophy of Science* 41: 188–95.

Tuomela, Raimo. 1980. Explaining Explaining. *Erkenntnis* 15: 211–43.

——. 1982. Explanation of Action. In *Contemporary Philosophy. A New Survey*, Vol. 3, ed. G. Floistad. The Hague, Boston and London: Martinus Nijhoff, 15–43.

Van Fraasen, Bas C. 1980. *The Scientific Image*. Oxford: Clarendon Press,

Wright, G. H. von. 1971. *Explanation and Understanding*. Ithaca, NY: Cornell University Press.