ALEXANDER BIRD

EXPLANATION AND METAPHYSICS

1. METAPHYSICS IN EXPLANATION

Is the nature of explanation a metaphysical issue? Or has it more to do with psychology and pragmatics? To put things in a different way: what are primary relata in an explanation? What sorts of thing explain what other sorts of thing? David Lewis identifies two senses of 'explanation' (Lewis 1986, 217–218). In the first sense, an explanation is an act of explaining. I shall call this the subjectivist sense, since its existence depends on some subject doing the explaining. Hence it is people who, in this sense, explain things. In the second of his two senses, Lewis says, quoting Sylvain Bromberger, that one may properly ask of an explanation "Does anyone know it? Who thought of it first? Is it very complicated?" (Lewis 1986, 218; Bromberger 1965). In this second sense, no subject is needed, the explanation can remain unknown, perhaps for ever. So I call this the objectivist sense.

Corresponding to the two senses, there are two tendencies, each emphasizing the primacy of one of the two senses. Corresponding to the subjectivist sense is an anti-metaphysical tendency that will typically deny that there are real explanations in nature. It is people who explain things; they do so by arranging information or beliefs in certain ways, so that certain salient connections may be seen. The saliency of such connections, and so whether an explanation is a good one or not, will depend on the interests of the recipient of the explanation. It cannot be denied that we do talk of one *fact* explaining another, as when we say that the decaying of the O-rings explained the explosion of the Challenger spacecraft. But this is just a projection of our practice of explanation onto the things we refer to in our explanations. Which is not to say that explanation is independent of the way things are; rather it is to say that it does not *constitute* part of the way things are. In this regard explanations are akin to secondary qualities, on a projectivist view of the latter.

Synthese (2005) 143: 89–107

The opposing tendency is to regard explanatory relations as part of the natural world, existing whether or not we do, waiting to be discovered along with other facts. For sure, we talk of people explaining things, but, in a reversal of the previous tendency, this talk is parasitic on the primary idea of facts explaining one another. When people explain things they describe possible explanations (in the primary sense) of facts of interest to them. Lewis, who thinks explanations in this sense are propositions, says it is unclear whether an incorrect explanation is just a bad explanation or no explanation at all (Lewis 1986, 218). But I suggest that this thought is perhaps a hangover from the subjectivist sense. Objectivist explanation (henceforth O-explanation) ought to be factive, for only if it is so does the issue of knowledge versus ignorance of explanations arise. We sometimes talk of the *real* explanation, to emphasize the factiveness of what we want to refer to. We may talk also, by contrast, of a proposed or putative explanation, in which case we are saying that it is proposed that (the fact) F is the explanation. Since 'F' occurs within an intensional context talk of 'putative explanation' is perfectly consistent with the claim that all explanations are facts. There might in this case be no fact F, and so the putative explanation is false – and hence no explanation at all.

Positivists and other empiricists typically exemplify the anti-metaphysical, 'subjectivist' tendency, while scientific realists are more likely to tend toward the objective, 'explanations-as-natural-phenomena' approach. Thus on the one side Comte rejected explanation altogether as being part of a disreputable metaphysics (Comte 1913), while van Fraassen gives a pragmatic account of explanation (van Fraassen 1980). On the other side, those who feel that inference to the best explanation, such as Lipton, Boyd, and Harman, is an indispensable part of our epistemic apparatus, will want to regard explanations as real features of the world, lest the central element of scientific inference be too subjective to be likely to lead to the truth (Lipton 1991; Boyd 1984; Harman 1965).

Now, it is true that this division is a rough one. For it is possible to be a logical empiricist, as was Hempel, yet give a metaphysical account of explanation – an account in terms of the actual laws and conditions obtaining (Hempel 1965). Nonetheless, it is significant that as his account came under increasing pressure, Hempel adopted a more and more subjectivist approach (Salmon 1990). At the same time, without wishing to accuse Halonen and Hintikka of anti-realism, it is also true that their account exhibits features of the anti-metaphysical approach. Hence there is a rejection of inference to the best explanation. Explanation is to be understood not in relation to laws, which are factual, but theories, which need not be. In this

response to their paper I shall try to promote a view of explanation with more metaphysical content that than provided by Halonen and Hintikka

But that is not to say that I disagree with the views they express – indeed I hope to show how a metaphysical account of objectivist explanation can complement their interrogative, pragmatic, view of subjectivist explanation. On the view I shall express both senses of explanation are legitimate. A central aim of subjectivist explanation (henceforth S-explanation) should be to describe parts of actual O-explanations. Which O-explanations and which parts of them are aimed at will be determined by our interests, including, as shall be explained, our epistemological interests. S-explanations may be judged as more or less *good*. What precisely goes into making an S-explanation good is of course a matter of philosophical debate. The sorts of relevant quality are: consistency with known facts (this is a minimum condition); consistency with propositions known to be likely to be true; simplicity; integration with other good explanations; power (i.e., that the S-explanation S-explains many facts). An epistemological question is whether these good-making features of an S-explanation relate to its describing an actual O-explanation, which is the central question in the issue of inference to the best explanation. This can be seen as a particular case of the relation between judgment (or assertion) and fact. Regarding judgment as an act, it aims (at least) at being true, i.e., describing some fact, just as in particular S-explanation aims at describing an O-explanation. Epistemology asks, what features of a judgment (or, in particular, an S-explanation) ensure or make it likely that it succeeds in describing a fact (or, correspondingly, an O-explanation)?

2. LAWS AND THEORIES

In this section I wish to illuminate some of the contrasts (and similarities) between Hempel's account of explanation and that provided by Halonen and Hintikka. The first contrast is one mentioned just a couple of paragraphs back, that they are offering an account of S-explanation, while Hempel is best understood as offering a model of O-explanation. Thus where Hempel has laws (whose existence is independent of us), Halonen and Hintikka have theories, which do depend on us. They tie explanation closely to epistemological issues (the interrogative model of inquiry), while epistemological issues are largely absent from Hempel's discussion.

But since, if I am right, there are two senses of 'explanation' at work here, it is not immediately clear that such contrasts point to a disagreement. So it may remain possible for us at one and the same time to accept Halonen and Hintikka's account as an account of S-explanation, while tak-

ing Hempel to have given us (at least the starting point for) an explication of O-explanation.

For example, Halonen and Hintikka say "By construing explanations as derivations from a background theory we can also dismiss one of the alleged difficulties with Hempel's model – or at least push it to a place difference from the theory of explanation. If explanation means subsumption under a scientific law, the next question will of course be: What is a scientific law? But if explaining means deriving the explanandum from a background theory, what an explainer has to do is to know which theories can be so used. And the obvious non-answer to that question is: The best available ones" (Halonen and Hintikka 2005). Instead of pushing the nature of laws out of the theory of explanation altogether, it is instead pushed out of the theory of S-explanation and into the metaphysical theory of O-explanation. If Halonen and Hintikka's theory of S-explanation and Hempel's theory of O-explanation are both right, then we should expect the background theory to state, if true, some actual laws of nature. What they say suggests that this is right "... the overall explanatory enterprise ... may very well involve the discovery of a suitable complex of background laws" (Halonen and Hintikka 2005). And, it should be added, in other cases of S-explanation, the laws are already known or believed to be known.

Before going further, some remarks should be addressed to the issue of the form of laws employed in a Hempel style explanation. Hempel's illustrations of the covering law model often employ laws of the form 'it is a law that all Fs are Gs', and correspondingly the initial or observational conditions are of the form F**a** to explain an explanandum of the form G**a**; i.e., overall the explanation has the form of a syllogistic argument. Halonen and Hintikka use Hempel's term 'covering law model' to refer to explanations with this sort of structure, and 'deductive-nomological' explanation to refer to a more general kind of explanation whose laws need not have the form of an universal generalization of an implication. (Hempel himself used 'covering-law model' as a general term including both D-N explanations and Inductive-Statistical explanations (Hempel 1965; 345–346)). It is not clear whether Hempel focused on laws and explanations of the narrow, syllogistic kind because they are illustratively more simple or because he thought that all laws and explanations may be cast in this form. He does include Newton's law of gravitation as a law which can figure in an explanation. So functional laws are included; but even these may be cast as complex, artificial universal generalisations. Be that as it may, Halonen and Hintikka are clearly right that it would be a mistake to focus only on the syllogistic form of explanation, and so in what follows, even where examples are syllogistic, this will not be taken to be a charac-

teristic of all explanations. So while some illustrations of O-explanations may be simple, actual O-explanations may be very complicated. However, it should be added that one difference between O-explanations (as I understand them) and S-explanations lies in the relations between the explanandum and explanans. Since S-explanation is something we do, the relation, as Halonen and Hintikka say, is *syntactic*. That is we *deduce* the explanandum from the background theory plus other evidence (nature's answers). However, O-explanation is independent of us and so the appropriate relation between explanans and explanandum is that of *semantic* entailment (in fact it is *more* than that, as I shall explain below). (Partly for this reason I regard a '*Deductive*-Nomological' as an inappropriate name for an account of O-explanation – I prefer '*Nomological-Instantiation*' – see below.) In a classical three body problem, the O-explanation will be simple: Newton's laws of gravitation and motion plus the initial conditions entail the explanandum final positions. But in an S-explanation, even if those laws are the content of our background theory and nature has given us those initial conditions as her answers, as Halonen and Hintikka point out, it will still be difficult to deduce the final positions. The difference is perhaps most stark in those non-linear cases where one thinks that the laws and initial conditions do entail the outcome without our being able to properly match this in any explanation we give, since the sensitivity of the outcome to initial conditions is always greater than the precision of nature's answers to questions about those conditions.

3. LAWS AND INSTANTIATION

As is well known, Hempel's accounts of law suffer from a variety of counter-examples and objections. Hempel's necessary and sufficient condition for explanation is that the explanandum be deducible from the total explanans, where the explanans consist of both laws and specific facts (which can be thought of as observed initial conditions). The main objections are: *triviality* – Hempel's account allows for the 'explanation' of facts by themselves; *preemption* – a fact may be deductively subsumed under several independent laws, but only one of these may actually explain the fact; and *symmetry* – a law may state a symmetrical relation between certain facts, so that X may be deduced from the law plus Y, and Y from the law plus X, so that Hempel's model allows each fact to explain the other. I suspect that the reason why some have regarded these as significant objections while Halonen and Hintikka do not find them serious or interesting is that the former have been concentrating on O-explanation, while Halonen and Hintikka are offering an account of S-explanation.

Similar objections would rise to their account, should it be considered as concerned with O-explanation. Let us take the classic case of preemption, as described by Achinstein (Achinstein 1983). The explanandum is the death of Jones. Hempel's model allows the explanation to be the fact that Jones ate a pound of arsenic, from which, with a law stating that death swiftly follows ingesting such a quantity of arsenic, Jones' death may be deduced. Halonen and Hintikka will allow the same result, since the fact about Jones' eating the poison may be produced as a result of an interrogative inquiry, and from it, with a background theory about the fatal effects of arsenic, Jones' death may be deduced. Achinstein's objection is that the facts stated thus far are consistent with its being the case that Jones died by some other means before the poison killed him; for instance, by being run down by a bus, or, for that matter, by being poisoned with a faster acting and almost undetectable poison that acts in a manner different from the arsenic. Halonen and Hintikka are right that the arsenic explanation, which fits their model, will satisfy everyone as an S-explanation of Jones' death. But as Achinstein's example shows, that is no guarantee that it is also the O-explanation of the death.

Therefore those concerned with O-explanation must take the counterexamples to Hempel into account. I will suggest that these flaws arise because Hempel, as an empiricist, was unable or unwilling to provide *enough* metaphysics for his account to work. First, as suggested in the last section, deducibility, being a syntactic notion, is inappropriate for an account of O-explanation. I shall argue that deducibility is in any case metaphysically too weak a notion to support an account of explanation, and that several of Hempel's problems flow from this.

The triviality objection may be circumvented by tightening up the conditions on the explanandum. However, the real problem is that on a regularity account of law, the explanation of an instance by a law will amount to the explanandum explaining itself. According to the regularity view, 'it is a law that Fs are Gs' is equivalent to 'all Fs are Gs' (perhaps plus some further condition, such as systematic integration with other nomic regularities). The proposition '**a** is F and G, **b** is F and G, **c** is F and G, . . . , and nothing is F other than **a**, **b**, **c**, . . . ' entails 'all Fs are Gs' and so should have enough content to explain anything the latter explains. But the proposition '**a** is F and G, **b** is F and G, **c** is F and G, . . . , and nothing is F other than **a**, **b**, **c**, . . . ' cannot explain an instance, viz. why **a**, which is F, is also G, because it just *states* that fact, plus facts about other entities, **b**, **c**, etc. that have no bearing on **a** itself. Hence 'all Fs are Gs' cannot explain an instance of that generalization. And so the regularity view fails to allow for laws to explain their instances (Bird 2002). Deductive subsumption

under a generalization is not O-explanation (Bird 1999). Hence Halonen and Hintikka are right to disqualify mere empirical generalizations from background theories in S-explanations (Halonen and Hintikka 2005).

The preemption and symmetry problems arise when we consider that deducing a proposition *p* from certain facts allows the truth of *p* to be predicted, but that is no guarantee of an explanation if *p* may also be deduced from some independent set of facts. This suggests that deductive subsumption of a fact under a law is no guarantee that the fact is a (or part of a) proper *instance* of a law. Here we may think of facts as properly instantiating the law when the exemplify the law *in action*. Some sets of facts may mimic the operation of an actual law, when the antecedent condition occurs (Jones ingests arsenic) and the appropriate consequent occurs (Jones dies), but this law has nothing to do with the matter – some other law or set of laws is responsible.

Thus, overall, we need notions of law and instantiation of a law so that:

(i) a law in not a mere regularity; there is some sort of ontological gap between a law and the sum of its instances;
(ii) the notion of a fact being a proper nomic instance of a law is stronger than the deducibility (or entailment) of the fact from the law.

The first desideratum may be achieved, for example, by employing a concept of law based upon nomic necessitation (Armstrong 1983; Tooley 1977; Dretske 1977), and some notion of instantiating a law which goes beyond deducibility. We may then spell out the framework of what one might call a *nomic-instantiation* (N-I) model of explanation thus: in outline, if N(F, G) is a law (which is a second order relation of necessitation, N, among the universals, F and G) and G**a** is the explanandum, then N(F, G) and F**a** together constitute the explanation of G**a** when ⟨F**a**, G**a**⟩ instantiates the law N(F, G). Here (a) ⟨X, Y⟩ instantiates the law L, entails (b) L&X entails Y, but (b) does not entail (a).

To flesh out the framework, it is required to give the notion of proper instantiation more content, satisfying condition (ii) above. Ultimately I do not believe that this notion can be spelled out with informative necessary and sufficient conditions. Necessarily not all concepts can. Nonetheless, one can draw some of its connections with other concepts and detail some of the conditions governing its use. Some, indeed most of the laws we come across, are not fundamental. In Armstrong's terms they are *derived* (Armstrong 1983). That is, they supervene on lower level laws. Thus law X is derived from the laws $L_1, \ldots L_n$ iff in all possible worlds in which $L_1, \ldots L_n$ are laws, X is also a law. Thus whenever X is 'in action' we should expect some appropriate non-empty subset of $L_1, \ldots L_n$ to be in action also. Above I contrasted Hempel's deductive subsumption with proper

instantiation. Let L be derived from a set of laws $L_1, \ldots L_n$ such that the subsumption (by entailment) of any set of facts under $L_1, \ldots L_n$ entails the existence of some fact that may be subsumed under L. Call such a set of laws a *sufficient subvenient* set for L. One condition we may formulate on proper instantiation is the following:

> ⟨X, Y⟩ properly instantiates the law L *only* if there is some sufficient subvenient set of *basic* laws for L and some set of facts that are subsumed under that set.

Thus subsumption, in something like the manner of Hempel is necessary for explanation, but not sufficient, for a further necessary condition is that deductive subsumption 'goes all the way down'. I do not think that the stronger condition is sufficient, since the issues of preemption and symmetry cannot be eliminated this way at the level of basic laws (****).

4. ELLIPTICAL EXPLANATION

It is quite clear, as Halonen and Hintikka emphasize, that the explanations we actually give are rarely anything like the explanations depicted by Hempel's model – or, for that matter, by my N-I account above. This however is not a weakness of either account, since both aim at explicating the idea of O-explanation. They do not aim at spelling out what happens in S-explanation. Nonetheless, there should be some connection between the objectivist and the subjectivist. As mentioned, on this view, an S-explanation aims at being or stating an O-explanation. A successful S-explanation will give the facts present in an O-explanation. However, an S-explanation need not typically aim at being a *complete* O-explanation. Rather it will aim at providing some of the facts contained in a complete O-explanation; which facts it aims at will be determined by pragmatic issues. In looking at the explanation of an accident we may be interested in some of the initial conditions, not in the laws (since we can change the former but not the latter). On the other hand, if we are interested in making scientific generalisations from evidence, we may be more interested in the laws than in the initial conditions. Such S-explanations, which focus on pragmatically determined elements in a complete O-explanation, are *elliptical* subjectivist explanations. Thus, when we focus on why questions, as Halonen and Hintikka enjoin us to do, the appropriate answer will be an elliptical explanation whose appropriateness will be determined by the pragmatic intent behind the why question. In such cases S-explanations are elliptical because out interests allow us to ignore some facts and encourage

us to focus on others. On other occasions the reason for providing an elliptical explanation is that we are ignorant of the complete explanation. But ignorance of the complete explanation is compatible with knowing some parts of it, and in particular the pragmatically interesting parts of it. (We may of course be in the less fortunate position of knowing some of the pragmatically uninteresting parts and being in ignorance of the parts we do care about.)

It is reasonable to identify causes of an explanandum fact with the initial conditions (the laws are not themselves causes). I will not dwell here on whether this is an entirely satisfactory way of understanding causal explanation. More important is to point out that in most complete O-explanations of a fact, there will be a large number of 'causes', i.e., initial conditions. Not all of these will be of relevant pragmatic interest. So, the accident investigators will not be interested, as mentioned, in the standing laws of nature. Nor will they be interested in other standing initial conditions (e.g., the fact that air contains oxygen is not a key fact of interest in explaining an engine fire, even if an initial condition of that fire); typically they will be interested in fact which do not obtain when the accident does not occur, e.g., a faulty electrical switch that sparks.

The fact that S-explanations are typically incomplete when compared to O-explanations, means that rarely will the information given in an S-explanation, even with all the background theory available, be enough to allow us to deduce the explanandum, even though this is what Halonen and Hintikka's Interrogative model requires. To be sure, they replace Hempel's deduction with questioning, but the purpose of the questioning is to get answers that will permit deduction of the explanandum from the background theory. Doctors may explain a patient's rare disease by pointing to a viral infection without hoping to be able to deduce the illness from that fact. The infection may be a necessary but not a sufficient condition of the illness, and background theory may not have enough detail to allow deduction of the illness, however many answers we get to questions put to nature. Similarly, the accident investigators may have very little in the way of background theory, not enough to deduce the fire from the explanans, but enough to show that the faulty switch is of explanatory significance.

Another kind of case where it will not be possible to deduce the explanandum from background theory plus nature's answers may arise when the background theory states laws with *ceteris paribus* clauses. In a poisoning case we may explain S's death by reference to the discovery that S ingested 30 g of arsenic and the background theory that arsenic is fatally poisonous in that quantity. But the claim that 30 g of arsenic is a fatal dose is one that must be hedged with 'other things being equal'. If S had

taken an antidote, or if S had immunised himself against arsenic poisoning by taking small, non-fatal doses over a long period, or if S had an very large body weight, or if S had an unusually robust constitution, then other things would not be equal or normal. Therefore, in order for S's death to be deducible from the background theory, we must put to nature the question 'are other things equal?' But that it just the sort of question nature cannot straightforwardly answer. We may put questions concerning the matters just listed, but the nature of a ceteris paribus law is just that the theory is unable to provide a conclusive list of all the factors to be excluded in order for conditions to be normal. In the given example our best evidence for the equality of other things is precisely that the poison did kill S. But it would clearly be circular then to use that proposition in deducing the death of S from the law of poisoning. The status of the proposition that other things are equal or normal is not that of an answer to a question put to nature, but that of a hypothesis, which remains to be confirmed. S-explanations employing ceteris paribus laws are also therefore inevitably epistemically incomplete.

Could the Interrogative model make do without the claim that the explanandum should be deducible from background theory and Nature's answers? Consider again the fire investigators. There will be an O-explanation for their case (indeed more than one, since there will be many mutually consistent explanations that differ in quantity of detail and historical scope). As remarked, the investigators' pragmatic interests direct them away from the unalterable laws of nature and other standing conditions. They are interested only in factors that are controllable and in particular factors that differ between the case of the fire and cases where there is no fire. So their interests force them to seek information about only a part of the O-explanation. And it is to this part that the interrogation of Nature will be directed. Nature's answer, along with background theory, will in such cases inevitably be insufficient to deduce propositions such as 'there is a fire'. Instead the answer should be a proposition p such that it is true that the facts (known or unknown) common to both the fire and the non-fire cases, together with p entail the existence of a fire; or, alternatively, such that, were it not for p there would have been no fire.

5. CONTRASTIVE EXPLANATION

This is why, as Halonen and Hintikka point out, many 'why' questions are implicitly or explicitly contrastive. This is not asking for a different sort of explanation. Rather the contrast is directing us towards some relevant part

of the complete explanation. Let us consider Halonen and Hintikka's set of why questions:

> Why did Tom fly to New York on Monday?
>
> Why did *Tom* fly to New York on Monday? (and not Dick or Harry)
>
> Why did Tom *fly* to New York on Monday? (instead of taking the train)
>
> Why did Tom fly to *New York* on Monday? (and not to Washington DC or Chicago)
>
> Why did Tom fly to New York on *Monday*? (and not Tuesday or Wednesday)

Lewis proposes that a question of the form "Why P rather than Q (or R or . . .)?" looks for "information about the causal history of the explanandum event [i.e., P], not including information that would also have applied to the causal histories of alternative events of the sort indicated [i.e., Q, R . . .], if one of them had taken place instead" (Lewis 1986, 229). That is, what is wanted is a cause of P that would not have been a cause of Q, had Q occurred. Say, for instance, it is Tom who flew to New York, not Dick. And the reason is that Tom was attending a high level meeting which required the attendance CEO of their firm, who is Tom. The fact that *Tom* was the CEO would not have played any part in causing Dick to go, had Dick gone to New York. That remains the case, even if it is the case that had Dick gone to New York the causal history of his doing so would have involved *his* being CEO (the nearest possible world in which Dick goes to New York is one in which Tom resigns and passes the job of CEO to Dick). Then it would have been *Dick's* being CEO (not Tom's) which would have played a causal role in his going.

However, we can see several faults with Lewis' proposal. The facts of the case might be that Tom seriously considered asking Dick to come as well, but decided on balance against. In such circumstances, the nearest possible world in which Dick goes to New York is one where Tom made the decision to take Dick too. In which case, Tom's being CEO would have played a causal role in Dick's going to New York – it is only because Tom is CEO that he needs Dick to go and can invite or require Dick to go. Therefore, Lewis would exclude a fact which we do regard as a satisfactory answer to the question "Why did *Tom*, and not *Dick*, fly to New York on Monday?", viz. Tom is CEO, might well have been a cause of Dick's going had Dick gone.

Hence, Lewis' account is too strong, ruling out some good contrastive explanations. More importantly perhaps, Lipton points out that Lewis'

account is also too weak (Lipton 1991, 42). Suppose that the meeting required only that some senior representative be present. In which case we might think that Tom's being a senior representative would not be a good explanation of why he went and not Dick, when Dick is also senior. But, had Dick gone, the fact that *Tom* is a senior representative would not have caused him to go (we imagine a case where Tom does not go because of illness). So Lewis' proposal allows the explanation through. If we broaden the notion of cause in this context to be a causal *type*, then we have problems of the first sort. Returning to the case where Tom goes because he is CEO, it might be true that had Dick gone it would have been because he, Dick, were CEO, hence ruling out a good explanation.

Lipton's approach is different. A necessary requirement on contrastive explanation is the 'Difference Condition': "to explain why P rather than Q we must cite a causal difference between P and not-Q, consisting of a cause of P and the absence of a corresponding event in the history of not-Q" (Lipton 1991, 43). In the first case considered, Tom's being CEO causes him to go, while Dick's being CEO is an event (or fact) which is absent, i.e., does not obtain, in the history of Dick's not going to New York. In the second case, the fact that Tom is a senior representative does not explain his going to New York and not Dick, since the Difference Condition is not met. The corresponding fact, Dick's being a senior representative, does obtain in his history.

The cases we have considered are *compatible* contrasts. When asking Why P rather than Q? in these cases, it is logically possible for both P and Q to occur. Both Tom and Dick might have gone to New York. But sometimes we ask contrastive questions when the P rules out Q-incompatible contrasts. So for instance, Lipton wants to explain why he went to the play *Jumpers* one night rather than *Candide*. He could not have gone to both. The problem here is that every event in the causal history of going to *Jumpers* is an event in the causal history of not going to *Candide*. So where do we find the difference? Lipton points out that the Difference Condition does not require that the explanation [of P rather than Q] not occur in the history of not-Q; rather the Condition requires only that there be no *corresponding* event in the history of not-Q. What would be the corresponding event? "... Roughly speaking, a corresponding event is something that would bear the same relation to Q as the cause of P bears to P" (Lipton 1991, 44). Lipton prefers contemporary plays and believes that *Jumpers* is a contemporary play. Had he gone to *Candide* it would have been because he (mistakenly) believed *Candide* to be a contemporary piece, or because his preference would have been for 18th century works. Both of these are absent in the causal history of not-Q.

Lipton says that his approach points to an actual difference between P and not-Q, rather than a counterfactual difference between P and Q, as Lewis would have it. But when we see the role played by the notion of a 'corresponding event' it begins to appear that this contrast is less sharp. Identifying the corresponding event requires us to consider the causal history of Q, had Q occurred, just as Lewis requires us to do. Putting the Difference Condition and the analysis of 'corresponding event' together, Lipton's account becomes: for E to be the explanation of why P rather than Q, E must be an event in the causal history of P such that had Q occurred, and an event E^* stood in the causal relation to Q that E stands in to P, then E^* is not to be found in the causal history of not-Q.

So Lewis is right that we look for a difference between the actual history of P and the counterfactual history of Q. But the difference is not something which is present in the former and absent in the latter. Rather, it is something which plays an analogous role in both and as it occurs in the one case (the history of P) is true of the actual world while as it occurs in the contrasting case (the history of Q) is false of the actual world.

However, some incompatible contrasts are a little trickier. Consider: why did Tom *fly* to New York, rather than take the train? Lewis considers a similar case which has the answer, there was too little time to go by train. How would Lipton's account handle this? We are supposed to look for the corresponding event, i.e., an event (or fact) which stands to Tom's going by train, had he gone by train, as his having too little time stands to his going by air. But nothing looks as if it does this. For the fact that there was too little time serves as a sufficient condition of his flying – it requires him to go by plane if he is to arrive before the meeting. Had he gone by train there would have been no event such that it required him to go by train for a timely arrival (after all he could still have gone by plane and arrived on time). It is true that had Tom gone by train, there would have been enough time for him to get to his meeting, but clearly the time requirement plays a different role with respect to his going by train. It is just a necessary condition, while the time consideration in the flying case is a sufficient condition. Nor does it help to shift to consideration of Tom's belief states (as we have in Lipton's theatre-going case). For it is not true that had Tom gone by train, he would have believed that only the train could get him there on time – that would be too far distant a possible world for this counterfactual.

So in requiring us to look in every case for a corresponding event is too strong. We should allow that it is sufficient for the Difference Condition to be met that there be no corresponding event – which we can regard as the trivial case of satisfying that Condition if it is expressed thus: it is not

the case that there is a corresponding event true of the causal history of not-Q. In effect, Lewis has taken his cue from considering the sorts of case where no corresponding event exists even in the would-be history of Q, and Lipton has focused on the cases where there is such an event.

Lipton contrasts his account with that of the D-N model. How do those comparisons affect the N-I approach I advocate? Lipton's first comment is that his analysis avoids the counterexamples to the D-N model-but then so does the I-N approach, except in certain basic cases where the N-I approach faces difficulties which also face any independent account of causation – which will therefore be problems for a causal analysis such as Lipton's. Secondly, says Lipton, his analysis "avoids the unhappy consequence that almost every explanation given is a mere sketch;; (Lipton 1991; 50) But it is not clear that this is such an unhappy consequence. Almost every S-explanation is elliptical and so omits to mention facts that obtain in a complete O-explanation. But that is unhappy only if the S-explanation aims at representing the complete O-explanation. Which typically it does not and should not; just as any representation, such as a drawing, need not aim at depicting every tiny detail of what is being represented (and may do so without being a *mere* sketch).

Therefore, the appropriate approach, as suggested above, is to regard contrastive why questions ("Why P rather than Q?") as directing the sketch – spelling out which facts from the complete O-explanation of P are worth mentioning in the elliptical S-explanation of P and which may be left out. Lipton's Difference Condition can be employed here. Facts to be left out include all those elements in the O-explanation of P that have counterparts in the counterfactual O-explanation of Q that exist in the actual world. That is, we are interested in facts in the explanation of P that do not have counterparts in the explanation of Q (Lewis' cases), or those that do have counterparts in the explanation of Q that do not exist in the actual world (Lipton's cases).

So a Lipton case looks like this. We wish to explain why Mervyn died after ingesting a pound of arsenic, while Mavis did not – the answer being that Mavis took an antidote but Mervyn did not. So we compare the actual explanation of Mervyn's dying and what would have been the explanation of Mavis dying, had she died. Using the D-N form for clarity (the same can be done with the N-I form):

Law	Everyone who ingests a pound of arsenic without an antidote dies within 24 hrs	Everyone who ingests a pound of arsenic without an antidote dies within 24 hrs

Conditions	Mervyn ingested a pound of arsenic Mervyn did not take an antidote (#)	Mavis ingested a pound of arsenic Mavis did not take an antidote (*)
entail	———————	———————
Explanandum	Mervyn died within 24hrs	Mavis died within 24 hrs

Here the elements in the O-explanation of Mervyn's death are, since O-explanations are factive, are all true in the actual world. The right hand explanation is a counterfactual O-explanation, true in a nearby possible world where Mavis does die. Some elements of this counterfactual are also true of the actual world (the first two lines). But the line marked (*), which states that Mavis did not take an antidote, is false in the actual world. So the line in the O-explanation of Mervyn's death in which we are interested, is the one which corresponds to (*), i.e., (#). Thus the explanation of the contrast is the fact that Mervyn did not take an antidote.

Now consider a Lewis case, explaining why Tom flew rather than taking the train:

Law	Tom does what maximizes his utility.	Tom does what maximizes his utility.
Conditions	Tom attaches high utility to his getting to New York before 11 am	Tom attaches high utility to his getting to New York before 11 am.
	Tom believes that only the plane will get him to NY before 11 am (#)	Tom believes that both the plane and train will get him to NY before 11 am
		Tom attaches a moderate marginal utility to the comfort of train travel over other forms of transport.
entail	———————	———————
Explanandum	Tom goes by plane	Tom goes by train

Here the detail of the structures of the two explanations is different, although there is some similarity. There is nothing on the right hand side which plays the same role as the fact (#) that Tom believes that only the plane will get him to NY before 11 am. So it will serve in a contrastive explanation, while the other facts on the left hand side do figure on the right hand side in a structurally similar role and are true in the actual world, and will thus be omitted. In general the answer to "Why P rather than Q?" will be obtained by constructing the explanation of P and the counterfactual explanation of Q, had Q occurred, and then looking for elements in the former that do not have a true counterpart in the latter.

One benefit of the D-N and N-I approaches to explanation is that they make easy sense of cases where what the contrastive why-question seeks is itself a law of nature. Why did the lump of lithium react when placed

in the beaker of water, but not the lump of boron? Here we may expect information concerning the laws of chemistry, laws governing the behaviour of lithium and contrasting laws concerning boron. While we are still dealing with causal explanation, it would be misleading to suggest that we are seeking differences in causal *histories*. Rather, it is contrasts in the laws being instantiated that are sought. As presented here, the Interrogative model suffers from this disadvantage, "... the background theory is taken for granted. A search for an explanation consists in an attempt to find the local facts ("initial conditions"). That together with the background theory imply the explanandum" (Halonen and Hintikka 2005). This requires that the background theory exist before the why-question is put, and the answer involves the local facts gained from interrogating nature. But sometimes what we need in order to explain the explanandum is not Nature's answers but a theory which does not exist yet. So when we ask, why do the Balmer lines exist? the answer does not employ a background theory plus Nature's answers to questions put thereafter. Rather the answer is a new theory, which does nor require any more of Nature's answers than were already available.

6. EXPLANATION AND THE INTERROGATIVE MODEL

I started by contrasting O-explanation, dealing with the way laws and facts relate metaphysically, and S-explanation, which is what is provided by an act of explanation. I said that Hempel's account aimed at the former, while I regarded the account provided by Hintikka's interrogative model as seeking to explicate the latter. This explains some of the divergences between the accounts considered, such as the reference to laws by Hempel and theories by Halonen and Hintikka. Such divergences are compatible with both being satisfactory explications of the respective kinds of explanation. Nonetheless, as stated, there is a link in that S-explanations seeks to state salient parts of actual O-explanations. Therefore there must be some correspondence between the two. I have suggested some ways in which Hempel's account of O-explanation might be improved. Is there anything we may learn about the interrogative model of inquiry as applied to S-explanation?

First, the requirement that the explanandum be deducible from background theory plus nature's answers is too strong. We may just know too little for this to be possible. But that should not prevent us from providing an S-explanation. It may be perfectly reasonable to suppose that unknown facts are such that with the facts averted to in the background theory and given in nature's answers, they all together entail the explanandum. For

instance, when our background theory contains *ceteris paribus* laws, we may be entitled to hypothesize that other things are equal, even if this is not known as an answer to a question put to nature.

This sort of elliptical S-explanation is called for in answer to contrastive why-questions In asking 'why P rather than Q?' the S-explanation aims at providing information about the O-explanation of P that has no true counterpart in the counterfactual O-explanation of Q. Therefore various standing conditions of which we may be ignorant, that are common to both O-explanations (including the fact that such conditions are normal or 'equal') will properly be omitted from the desired contrastive S-explanation.

Secondly, it should be borne in mind that even a complete S-explanation whose elements (background theory, nature's answers and so on) are all true, has no guarantee of corresponding to the O-explanation of the explanandum, as the objections to Hempel's model show. That, of course, is no criticism of a theory of S-explanation. It does show, however, that a complete S-explanation is like an elliptical or incomplete one in being essentially hypothetical. The issue of when an S-explanation can be regarded as giving the actual O-explanation is the key question of scientific epistemology. Standard accounts of Inference to the Best Explanation (Lipton 1991; Harman 1965) suggest that we select among given S-explanations according to criteria of goodness or 'loveliness', and the best S-explanation is most likely to be the actual O-explanation (or relevant part thereof). Another approach, which I prefer, seeks to eliminate all S-explanations but one. On this view an S-explanation may be eliminated if it requires the truth of some proposition known to be false. In general there may be several ways in which this may be done. First, and trivially, an S-explanation may be eliminated if any of its components are discovered to be false. Secondly, if the S-explanation is incomplete, even where its components are all true, there will be some conditions that would be required to make the explanation complete and of which we are currently ignorant. We may later come to discover that some such condition cannot be fulfilled. A doctor may suspect on the basis of certain symptoms that a patient has a certain unusual disease. But if it can be shown that the patient has never come into contact with a carrier of that disease we can eliminate that hypothesis and look for another, if the nature of the disease is such that infection from a carrier would have to be part of any complete possible O-explanation involving that disease. Thirdly, even if an S-explanation is complete, it may still entail some proposition we can discover to be false. This will be the case, for instance, if we find that an essential subvenient law is not instantiated. So for instance we may

eliminate the arsenic S-explanation of Jones' death if we discover that at the time of his death there was no significant organ failure, since arsenic as an O-explanation of death would require this. Hence the arsenic S-explanation may be eliminated. Thus a key role for the Interrogative model of enquiry is to ask questions of nature whose answers may serve to eliminate S-explanations. If only one S-explanation remains that is consistent with what we know from nature's answers (or what is elicited as having been tacitly known (Hintikka and Hintikka 1982)) then we may deduce that this S-explanation is true, i.e., states the actual O-explanation. This is deduced in accordance with Sherlock Holmes' famous dictum "Eliminate the impossible and whatever remains, however improbable, must be the truth". Thus understood, Inference to the Best Explanation (i) is deductive, and (ii) exemplifies the Interrogative model; this view of Inference to the Best Explanation is in conformity with and indeed vindicates Hintikka's views as expressed on information seeking by questioning (Hintikka and Hintikka 1982). Of course, there remains the matter of whether we can ever know enough to eliminate all but one of the S-explanations we should consider. But that is another question.

REFERENCES

Achinstein, P.: 1983, *The Nature of Explanation*, Oxford University Press, Oxford.
Armstrong, D.: 1983, *What is a Law of Nature?* Cambridge University Press, Cambridge.
Bird, A.: 1999, 'Explanation and Laws', *Synthese* **120**.
Bird, A.: 2002, 'Laws and Criteria', *Canadian Journal of Philosophy* **32**.
Boyd, R.: 1984, 'On the Current Status of Scientific Realism', in J. Leplin (ed.), *Scientific Realism*, University of California Press, Berkeley.
Bromberger, S.: 1965, 'An Approach to Explanation', in R. J. Butler (ed.), *Analytical Philosophy*, Second Series, Blackwell, Oxford.
Comte, A.: 1913, *The Positive Philosophy of Auguste Comte* (trans. H. Martineau), Chiswick, London.
Dretske, F.: 1977, 'Laws of Nature', *Philosophy of Science* **44**.
Halonen, I. and Hintikka, J.: 2005, 'Toward a Theory of the Process of Explanation', *Synthese* **143**, 5–61.
Harman, G.: 1965, 'The Inference to the Best Explanation', *Philosophical Review* **74**.
Hempel, C.: 1965, *Aspects of Scientific Explanation*, Free Press, New York.
Hintikka, J. and M. Hintikka: 1982, 'Sherlock Holmes Confronts Modern Logic: Towards a Theory of Information seeking by Questioning', in E. Barth and J. Martens (eds.), *Argumentation: Approaches to Theory Formation*, John Benjamins, Amsterdam.
Lewis, D.: 1986, 'Causal Explanation', in *Philosophical Papers* vol. II, Oxford University Press, Oxford.
Lipton, P.: 1991, *Inference to the Best Explanation*, Routledge, London.
Salmon, W.: 1990, *Four Decades of Scientific Explanation*, Princeton University Press, Princeton.

Tooley, M.: 1977, 'The Nature of Laws', *Canadian Journal of Philosophy* **7**.
Van Fraassen, B.: 1980, *The Scientific Image*, Oxford University Press, Oxford.

The University of Bristol
Department of Philosophy
9 Woodland Road
Bristol BS8 1TB, U.K.