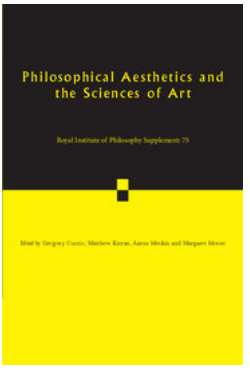

Contrastive Explanation

Peter Lipton

Royal Institute of Philosophy Supplement / Volume 27 / March 1990, pp 247 - 266
DOI: 10.1017/S1358246100005130, Published online: 08 January 2010

Contrastive Explanation*

PETER LIPTON

1. Introduction

According to a causal model of explanation, we explain phenomena by giving their causes or, where the phenomena are themselves causal regularities, we explain them by giving a mechanism linking cause and effect. If we explain why smoking causes cancer, we do not give the cause of this causal connection, but we do give the causal mechanism that makes it. The claim that to explain is to give a cause is not only natural and plausible, but it also avoids many of the objections to other accounts of explanation, such as the views that to explain is to give a reason to believe the phenomenon occurred, to somehow make the phenomenon familiar, or to give a Deductive-Nomological argument. Unlike the reason for belief account, a causal model makes a clear distinction between understanding why a phenomenon occurs and merely knowing that it does, and the model does so in a way that makes understanding unmysterious and objective. Understanding is not some sort of super-knowledge, but simply more knowledge: knowledge of the phenomenon and knowledge of its causal history. A causal model makes it clear how something can explain without itself being explained, and so avoids the regress of whys, since we can know a phenomenon's cause without knowing the cause of the cause. It also accounts for legitimate self-evidencing explanations, explanations where the phenomenon is an essential part of the evidence that the explanation is correct, so the explanation can not supply a non-circular reason for believing the phenomenon occurred. There is no barrier to knowing a cause through its effects and also knowing that it is their cause. The speed of recession of a star explains its observed red-shift, even though the shift is an essential part of the evidence for its speed of recession. The model also avoids the most serious objection to the familiarity view, which is that some phenomena are familiar yet not understood, since a phenomenon can be perfectly familiar, such as the blueness of the sky or the fact that the same side of the moon always

* I am grateful to Philip Clayton, Trevor Hussey, Philip Pettit, David Ruben, Elliot Sober, Edward Stein, Nick Thompson, Jonathan Vogel, David Weissbord, Alan White, Tim Williamson, and Eddy Zemach for helpful discussions about contrastive explanation.

faces the earth, even if we do not know its cause. Finally, a causal model avoids many of the objections to the Deductive-Nomological model. Ordinary explanations do not have to meet the requirements of the Deductive-Nomological model, because one does not need to give a law to give a cause, and one does not need to know a law to have good reason to believe that a cause is a cause. As for the notorious over-permissiveness of the Deductive-Nomological model, the reason recession explains red-shift but not conversely, is simply that causes explain effects but not conversely, and the reason a conjunction of laws does not explain its conjuncts is that conjunctions do not cause their conjuncts.

The most obvious objection to a causal model of explanation is that there are non-causal explanations. Mathematicians and philosophers give explanations, but mathematical explanations are never causal, and philosophical explanations seldom are. A mathematician may explain why Gödel's Theorem is true, and a philosopher may explain why there can be no inductive justification of induction, but these are not explanations that cite causes. In addition to the mathematical and philosophical cases, there are explanations of the physical world that seem non-causal. Here is a personal favourite. Suppose that some sticks are thrown into the air with a lot of 'spin', so that they separate and tumble about as they fall. Now freeze the scene at a certain point during the sticks' descent. Why are appreciably more of them near the vertical axis than near the horizontal, rather than in more or less equal numbers near each orientation, as one would have expected? The answer, roughly speaking, is that there are many more ways for a stick to be near the horizontal than near the vertical. To see this, consider purely horizontal and vertical orientations for a single stick with a fixed midpoint. There are infinitely many of the former, but only two of the latter. Less roughly, the explanation is that there are two horizontal dimensions but only one vertical one. This is a lovely explanation, but apparently not a causal one, since geometrical facts cannot be causes.

Non-causal explanations show that a causal model of explanation cannot be complete. Nevertheless, a causal model is still a good bet now, because of the backward state of alternate views of explanation, and the overwhelming preponderance of causal explanations among all explanations. Nor is it *ad hoc* to limit our attention to causal explanations. A causal model does not simply pick out a feature that certain explanations happen to have: causal explanations are explanatory *because* they are causal. Like other accounts of explanation, however, causal models face a problem of underdetermination. Most causes do not provide good explanations. This paper attempts a partial solution to this problem.

2. Fact and Foil

Let us focus on the causal explanation of particular events. The problem here is with the notion of explaining an event by giving *the* cause. We may explain an event by giving some information about its causal history (Lewis, 1986), but causal histories are long and wide, and most causal information does not provide a good explanation. The big bang is part of the causal history of every event, but explains only a few. The spark and the oxygen are both part of the causal history that led up to the fire, but only one of them explains it. So what makes one piece of information about the causal history of an event explanatory and another not? The short answer is that the cause that explains depends on our interests. But this does not yield a very informative model of explanation unless we can go some way towards spelling out how explanatory interests determine explanatory causes.

One way to show how we select from among causes is to reveal additional structure in the phenomenon to be explained, structure that points to a particular cause. We can account for the specificity of explanatory answers by revealing the specificity in the explanatory question. Suppose we started by construing a phenomenon to be explained simply as a concrete event, say a particular eclipse. The number of causal factors is enormous. As Hempel has observed, however, we do not explain events, only aspects of events (Hempel, 1965, pp. 421–3). We do not explain the eclipse *tout court*, but only why it lasted as long as it did, or why it was partial, or why it was not visible from a certain place. This reduces the number of causal factors we need consider for any particular phenomenon, since there will be many causes of the eclipse that are not, for example, causes of its duration.

More recently, it has been argued that explanation is 'interest relative', and that we can analyse some of this relativity with a contrastive analysis of the phenomenon to be explained. What gets explained is not simply 'Why this', but 'Why this rather than that' (Garfinkel, 1981, pp. 28–41; van Fraassen, 1980, pp. 126–9). A contrastive phenomenon consists of a *fact* and a *foil*, and the same fact may have several different foils. We may not explain why the leaves turn yellow in November *tout court*, but only, for example, why they turn yellow in November rather than in January, or why they turn yellow in November rather than turning blue.

The contrastive analysis of explanation is extremely natural. We often pose our why-questions in explicitly contrastive form and it is not difficult to come up with examples where different people select different foils, requiring different explanations. When I asked my three year old son why he threw his food on the floor, he told me that he was full. This may explain why he threw it on the floor rather than eating it,

but I wanted to know why he threw it rather than leaving it on his plate. Similarly, an explanation of why I went to see *Jumpers* rather than *Candide* will probably not explain why I went to see *Jumpers* rather than staying at home, and an explanation of why Able rather than Baker got the philosophy job may not explain why Able rather than Charles got the job. The proposal that phenomena to be explained have a complex fact–foil structure can be seen as another step along Hempel's path of focusing explanation by adding structure to the why-question. A fact is usually not specific enough: we also need to specify a foil. Since the causes that explain a fact relative to one foil will not generally explain it relative to another, the contrastive question provides a further restriction on explanatory causes.

While the role of contrasts in explanation will not account for all the factors that determine which cause is explanatory, I believe that it does provide the central mechanism. In this essay, I want to show in some detail how contrastive questions help select explanatory causes. My discussion will fall into three parts. First, I will make three general observations about contrastive explanation. Then, I will use these observations to show why contrastive questions resist reduction to non-contrastive form. Finally, I will describe the mechanism of 'causal triangulation' by which the choice of foils in contrastive questions helps to select explanatory causes.

When we ask a contrastive why-question—'Why the fact rather than the foil?'—we presuppose that the fact occurred and that the foil did not. Often we also suppose that the fact and the foil are in some sense incompatible. When we ask why Kate rather than Frank won the Philosophy Department Prize, we suppose that they could not both have won. Similarly, when we asked about leaves, we supposed that if they turn yellow in November, they cannot turn yellow in January, and if they turn yellow in November they cannot also turn blue then. Indeed, it is widely supposed that fact and foil are always incompatible (Garfinkel, 1981, p. 40; Ruben, 1987; Temple, 1988, p. 144). My first observation is that this is false: many contrasts are compatible. We often ask a contrastive question when we do not understand why two apparently similar situations turned out differently. In such a case, far from supposing any incompatibility between fact and foil, we ask the question just because we expected them to turn out the same. By the time we ask the question, we realize that our expectation was disappointed, but this does not normally lead us to believe that the fact precluded the foil, and the explanation for the contrast will usually not show that it did. Consider the much discussed example of syphilis and paresis (cf. Hempel, 1965, pp. 369–70; van Fraassen, 1980, p. 128). Few with syphilis contract paresis, but we can still explain why Jones rather than Smith contracted paresis by pointing out that only Jones

had syphilis. In this case, there is no incompatibility. Only Jones contracted paresis, but they both could have: Jones's affliction did not protect Smith. Of course not every pair of compatible propositions would make a sensible contrast but, as we will eventually see, it is not necessary to restrict contrastive questions to incompatible contrasts to distinguish sensible questions from silly ones.

My second and third observations concern the relationship between an explanation of the contrast between a fact and foil and the explanation of the fact alone. I do not have a general account of what it takes to explain a fact on its own. As we will see, this is not necessary to give an account of what it takes to explain a contrast; indeed, this is one of the advantages of a contrastive analysis. Yet, based on our intuitive judgments of what is and what is not an acceptable explanation of a fact alone, the requirements for explaining a fact diverge from the requirements for explaining a contrast. My second observation, then, is that explaining a contrast is sometimes easier than explaining the fact alone (cf. Garfinkel, 1981, p. 30). An explanation of 'P rather than Q' is not always an explanation of P. This is particularly clear in examples of compatible contrasts. Jones's syphilis does not explain why he got paresis, since the vast majority of people who get syphilis do not get paresis, but it does explain why Jones rather than Smith got paresis, since Smith did not have syphilis. The relative ease with which we explain some contrasts also applies to many cases where there is an incompatibility between fact and foil. My preference for contemporary plays may not explain why I went to see *Jumpers* last night, since it does not explain why I went out, but it does explain why I went to see *Jumpers* rather than *Candide*. A particularly striking example of the relative ease with which some contrasts can be explained is the explanation that I chose A rather than B because I did not realize that B was an option. If you ask me why I ordered eggplant rather than sea bass (a 'daily special'), I may give the perfectly good answer that I did not know there were any specials, but this would not be an acceptable answer to the simple question, 'why did you order eggplant?' One reason we can sometimes explain a contrast without explaining the fact alone seems to be that contrastive questions incorporate a presupposition that makes explanation easier. To explain 'P rather than Q' is give a certain type of explanation of P, *given* 'P or Q', and an explanation that succeeds with the presupposition will not generally succeed without it.

My final observation is that explaining a contrast is also sometimes harder than explaining the fact alone. An explanation of P is not always an explanation of 'P rather than Q'. This is obvious in the case of compatible contrasts: you cannot explain why Jones rather than Smith contracted paresis without saying something about Smith. But it also applies to incompatible contrasts. To explain why I went to *Jumpers*

rather than *Candide*, it is not enough for me to say that I was in the mood for a philosophical play. To explain why Kate rather than Frank won the prize, it is not enough that she wrote a good essay; it must have been better than Frank's. One reason that explaining a contrast is sometimes harder than explaining the fact alone is that explaining a contrast requires giving causal information that distinguishes the fact from the foil, and information that we accept as an explanation of the fact alone may not do this.

3. Failed Reductions

There have been a number of attempts to reduce contrastive questions to non-contrastive and generally truth-functional form. One motivation for this is to bring contrastive explanations into the fold of the Deductive-Nomological model since, without some reduction, it is not clear what the conclusion of a deductive explanation of 'P rather than Q' ought to be. Armed with our three observations—that contrasts may be compatible, and that explaining a contrast is sometimes easier and sometimes harder than explaining the fact alone—we can show that contrastive questions resist a reduction to non-contrastive form. We have already seen that the contrastive question 'Why P rather than Q?' is not equivalent to the simple question 'Why P?', where two why-questions are explanatorily equivalent just in case any adequate answer to one is an adequate answer to the other. One of the questions may be easier or harder to answer than the other. Still, a proponent of the Deductive-Nomological model of explanation may be tempted to say that, for incompatible contrasts, the question 'Why P rather than Q?' is equivalent to 'Why P?' But it is not plausible to say that a Deductive-Nomological explanation of P is generally necessary to explain 'P rather than Q'. And it is even dubious that a Deductive-Nomological explanation of P is always sufficient to explain 'P rather than Q'. Imagine a typical deductive explanation for the rise of mercury in a thermometer. Such an explanation would explain various contrasts, for example why the mercury rose rather than fell. It may not, however, explain why the mercury rose rather than breaking the glass. A full Deductive-Nomological explanation of the rise will have to include a premise saying that the glass does not break, but it does not need to explain this.

Another natural suggestion is that the contrastive question 'Why P rather than Q?' is equivalent to the conjunctive question 'Why P and not-Q?' On this view, explaining a contrast between fact and foil is tantamount to explaining the conjunction of the fact and the negation of the foil (Temple, 1988). In ordinary language, a contrastive question is often equivalent to its corresponding conjunction, simply because the

'and not' construction is often used contrastively. Instead of asking, 'Why was the prize won by Kate rather than by Frank?', the same question could be posed by asking 'Why was the prize won by Kate and not by Frank?'. But this colloquial equivalence does not seem to capture the point of the conjunctive view. So I suggest that the conjunctive view be taken to entail that explaining a conjunction at least requires explaining each conjunct; than an explanation of 'P and not-Q' must also provide an explanation of P and an explanation of not-Q. Thus, on the conjunctive view, to explain why Kate rather than Frank won the prize at least requires an explanation of why Kate won it and an explanation of why Frank did not. This account of contrastive explanation falls to the observation that explaining a contrast is sometimes easier than explaining the fact alone, since explaining P and explaining not-Q is at least as difficult as explaining P.

The observations that explaining contrasts is sometimes easier and sometimes harder than explaining the fact alone reveal another objection to the conjunctive view, on any model of explanation that is deductively closed. (A model is deductively closed if it entails that an explanation of P will also explain any logical consequence of P.) Consider cases where the fact is logically incompatible with the foil. Here P entails not-Q, so the conjunction 'P and not-Q' is logically equivalent to P alone. Furthermore, all conjunctions whose first conjunct is P and whose second conjunct is logically incompatible with P will be equivalent to each other, since they are all logically equivalent to P. Hence, for a deductively closed model of explanation, explaining 'P and not-Q' is tantamount to explaining P, whatever Q may be, so long as it is incompatible with P. We have seen, however, that explaining 'P rather than Q' is not generally tantamount to explaining P. The conjunction is explanatorily equivalent to P, and the contrast is not, so the conjunction is not equivalent to the contrast.

The failure to represent a contrastive phenomenon by the fact alone or by the conjunction of the fact and the negation of the foil suggests that, if we want a non-contrastive paraphrase, we ought instead to try something logically weaker than the fact. In some cases, it does seem that an explanation of the contrast is really an explanation of a logical consequence of the fact. This is closely related to what Hempel has to say about 'partial explanation' (1965, pp. 415–18). He gives the example of Freud's explanation of a particular slip of the pen that resulted in writing down the wrong date. Freud explains the slip with his theory of wish-fulfillment, but Hempel objects that the explanation does not really show why that particular slip took place, but at best only why there was some wish-fulfilling slip or other. Freud gave a partial explanation of the particular slip, since he gave a full explanation of the weaker claim that there was some slip. Hempel's point fits naturally

into contrastive language: Freud did not explain why it was this slip rather than another wish-fulfilling slip, though he did explain why it was this slip rather than no slip at all. And it seems natural to analyse 'Why this slip rather than no slip at all?' as 'Why some slip?'

In general, however, we cannot paraphrase contrastive questions with consequences of their facts. We cannot, for example, say that to explain why the leaves turn yellow in November rather than in January is just to explain why the leaves turn (some colour or other) in November. This attempted paraphrase fails to discriminate between the intended contrastive question and the question, 'Why do the leaves turn in November rather than falling right off?' Similarly, we cannot capture the question, 'Why did Jones rather than Smith get paresis?', by asking about some consequence of Jones's condition, such as why he contracted a disease.

A general problem with finding a paraphrase entailed by the fact P is that, as we have seen, explaining a contrast is sometimes harder than explaining P alone. There are also problems peculiar to the obvious candidates. The disjunction, 'P or Q' will not do: explaining why I went to *Jumpers* rather than *Candide* is not the same as explaining why I went to either. Indeed, this proposal gets things almost backwards: the disjunction is what the contrastive question assumes, not what calls for explanation. This suggests, instead, that the contrast is equivalent to the conditional, 'if P or Q, then P' or, what comes to the same thing if the conditional is truth-functional, to explaining P on the assumption of 'P or Q'. Of all the reductions we have considered, this proposal is the most promising, but I do not think it will do. On a deductive model of explanation it would entail that any explanation of not-Q is also an explanation of the contrast, which is incorrect. We cannot explain why Jones rather than Smith has paresis by explaining why Smith did not get it. It would also wrongly entail that any explanation of P is an explanation of the contrast, since P entails the conditional.

4. Causal Triangulation

By asking a contrastive question, we can achieve a specificity that we do not seem to be able to capture either with a non-contrastive sentence that entails the fact or with one that the fact entails. But how then does a contrastive question specify the sort of information that will provide an adequate answer? It now appears that looking for a non-contrastive reduction of 'P rather than Q' is not a useful way to proceed. The contrastive claim may entail no more than 'P and not-Q' or perhaps better, 'P but not-Q', but explaining the contrast is not the same as explaining these conjuncts. We will do better to leave the analysis of the

contrastive question to one side, and instead consider directly what it takes to provide an adequate answer. David Lewis has given an interesting account of contrastive explanation that does not depend on paraphrasing the contrastive question. According to him, we explain why event P occurred rather than event Q by giving information about the causal history of P that would not have applied to the history of Q, if Q had occurred (Lewis, 1986, pp. 229–30). Roughly, we cite a cause of P that would not have been a cause of Q. In Lewis's example, we can explain why he went to Monash rather than to Oxford in 1979 by pointing out that only Monash invited him, because the invitation to Monash was a cause of his trip, and that invitation would not have been a cause of a trip to Oxford, if he had taken one. On the other hand, Lewis's desire to go to places where he has good friends would not explain why he went to Monash rather than Oxford, since he has friends in both places and so the desire would have been part of either causal history.

Lewis's account, however, is too weak: it allows for unexplanatory causes. Suppose that both Oxford and Monash had invited him, but he went to Monash anyway. On Lewis's account, we can still explain this by pointing out that Monash invited him, since that invitation still would not have been a cause of a trip to Oxford. Yet the fact that he received an invitation from Monash clearly does not explain why he went there rather than to Oxford in this case, since Oxford invited him too. Similarly, Jones's syphilis satisfies Lewis's requirement even if Smith has syphilis too, yet in this case it would not explain why Jones rather than Smith contracted paresis.

It might be thought that Lewis's account could be saved by construing the causes more broadly, as types rather than tokens. In the case of the trip to Monash, we might take the cause to be receiving an invitation rather than the particular invitation to Monash he received. If we do this, we can correctly rule out the attempt to explain the trip by appeal to an invitation if Oxford also invited since, in this case, receiving an invitation would also have been a cause of going to Oxford. This, however, will not do, for two reasons. First, it does not capture Lewis's intent: he is interested in particular elements of a particular causal history, not general causal features. Secondly, and more importantly, the suggestion throws out the baby with the bath water. Now we have also ruled out the perfectly good explanation by invitation in some cases where only Monash invites. To see this, suppose that Lewis is the sort of person who only goes where he is invited. In this case, an invitation would have been part of a trip to Oxford, if he had gone there.

To improve on Lewis's account, consider John Stuart Mill's Method of Difference, his version of the controlled experiment (Mill, 1904, bk. III, ch. VIII, sec. 2). Mill's method rests on the principle that a cause

must lie among the antecedent differences between a case where the effect occurs and an otherwise similar case where it does not. The difference in effect points back to a difference that locates a cause. Thus we might infer that contracting syphilis is a cause of paresis, since it is one of the ways Smith and Jones differed. The cause that the Method of Difference isolates depends on which control we use. If, instead of Smith, we have Doe, who does not have paresis but did contract syphilis and had it treated, we would be led to say that a cause of paresis is not syphilis, but the failure to treat it. The Method of Difference also applies to incompatible as well as to compatible contrasts. As Mill observes, the method often works particularly well with diachronic (before and after) contrasts, since these give us histories of fact and foil that are largely shared, making it easier to isolate a difference. If we want to determine the cause of a person's death, we naturally ask why he died when he did rather than at another time, and this yields an incompatible contrast, since you can only die once.

The Method of Difference concerns the discovery of causes rather than the explanation of effects, but the similarity to contrastive explanation is striking (cf. Garfinkel, 1981, p. 40). Accordingly, I propose that, for the causal explanations of events, explanatory contrasts select causes by means of what I will call 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.* Instead of pointing to a counterfactual difference, a particular cause of P that would not have been a cause of Q, as Lewis suggests, contrastive questions select as explanatory an actual difference between P and not-Q. Lewis's invitation to Monash does not explain why he went there rather than to Oxford if he was invited to both places because, while there is an invitation in the history of his trip to Monash, there is also an invitation in the history that leads him to forgo Oxford. Similarly, the Difference Condition correctly entails that Jones's syphilis does not explain why he rather than Smith contracted paresis if Smith had syphilis too, and that Kate's submitting an essay does not explain why she rather than Frank won the prize. Consider now some of the examples of successful contrastive explanation. If only Jones had syphilis, that explains why he rather than Smith has paresis, since having syphilis is a condition whose presence was a cause of Jones's paresis and a condition that does not appear in Smith's medical history. Writing the best essay explains why Kate rather than Frank won the prize, since that marks a causal difference between the two of them. Lastly, the fact that *Jumpers* is a contemporary play and *Candide* is not caused me both to go to one and to avoid the other.

The application of the Difference Condition is easiest to see in cases of compatible contrasts, since here the causal histories of P and of not-Q are generally distinct, but the condition does not require this. In cases of choice, for example, the causal histories are usually the same: the causes of my going to *Jumpers* are the same as the causes of my not going to *Candide*. The Difference Condition may nevertheless be satisfied if my belief that *Jumpers* is a contemporary play is a cause of going, and I do not believe that *Candide* is a contemporary play. That is why my preference for contemporary plays explains my choice. The Difference Condition does not require that the same event be present in the history of P but absent in the history of not-Q, a condition that could never be satisfied when the two histories are the same, but only that the cited cause of P find no corresponding event in the history of not-Q, where a corresponding event is something that would bear the same relation to Q as the cause of P bears to P.

One of the merits of the Difference Condition is that it brings out the way the incompatibility of fact and foil, when it obtains, is not sufficient to transform an explanation of the fact into an explanation of the contrast, even if the cause of the fact is also a cause of the foil not obtaining. Perhaps we could explain why Able got the philosophy job by pointing out that Quine wrote him a strong letter of recommendation, but this will only explain why Able rather than Baker got the job if Quine did not also write a similar letter for Baker. If he did, Quine's letter for Able does not alone explain the contrast, even though that letter is a cause of both Able's success and Baker's failure, and the former entails the latter. The letter may be a partial explanation of why Able got the job, but it does not explain why Able rather than Baker got the job. In the case where they both have strong letters from Quine, a good explanation of the contrast will have to find an actual difference, say that Baker's dossier was weaker than Able's in some other respect, or that Able's specialities were more useful to the department. There are some cases of contrastive explanation that do seem to rely on the way the fact precludes the foil, but I think these can be handled by the Difference Condition. For example, suppose we explain why a bomb went off prematurely at noon rather than in the evening by saying that the door hooked up to the trigger was opened at noon (I owe this example to Eddy Zemach). Here it may appear that the Difference Condition is not in play, since the explanation would stand even if the door was also opened in the evening. But the Difference Condition is met, if we take the cause not simply to be the opening of the door, but the opening of the door when it is rigged to an armed bomb.

My goal in this paper is to show how the choice of contrast helps to determine an explanatory cause, not to show why we choose one contrast rather than another. Still, some account of the considerations

that govern our choice of why-questions would have to be a part of a full model of our explanatory practices, and it is to the credit of my model of contrastive explanation that it lends itself to this. For example, as I have already observed, not all contrasts make for sensible contrastive questions. It does not make sense, for example, to ask why Lewis went to Monash rather than Baker getting the philosophy job. One might have thought that a sensible contrast must be one where fact and foil are incompatible, but we have seen that this is not necessary, since there are many sensible compatible contrasts. There are also incompatible contrasts that do not yield reasonable contrastive questions, such as why someone died when she did rather than never having been born. The Difference Condition suggests instead that the central requirement for a sensible contrastive question is that the fact and the foil have a largely similar history, against which the differences stand out. When the histories are disparate, we do not know where to begin to answer the question. There are, of course, other considerations that help to determine the contrasts we actually choose. For example, in the case of incompatible contrasts, we often pick as foil the outcome we expected; in the case of compatible contrasts, as I have already mentioned, we often pick as foil a case we expected to turn out the same way as the fact. The condition of a similar history also helps to determine what will count as a corresponding event. If we were to ask why Lewis went to Monash rather than Baker getting the job, it would be difficult to see what in the history of Baker's failure would correspond to Lewis's invitation, but when we ask why Able rather than Baker got the job, the notion of a corresponding event is relatively clear.

5. Further Issues

I will now consider three further issues connected with my analysis of contrastive explanation: the need for further principles for distinguishing explanatory from unexplanatory causes, the prospects for treating all why-questions as contrastive, and a comparison of my analysis with the Deductive-Nomological model. When we ask contrastive why-questions, we choose our foils to point towards the sort of causes that interest us. As we have just seen, when we ask about a surprising event, we often make the foil the thing we expected. Failed expectations are not, however, the only things that prompt us to ask why-questions. If a doctor is interested in the internal etiology of a disease, he will ask why the afflicted have it rather than other people in similar circumstances, even though the shared circumstances may be causally relevant to the disease. Again, if a machine malfunctions, the natural diagnostic contrast is its correct behaviour, since that directs our attention to the

causes that we want to change. But the contrasts we construct will almost always leave multiple differences that meet the Difference Condition, and this raises the problem of selecting from among them. A problem of multiple differences also arises for the Method of Difference, in the context of inference rather than explanation. Mill tells us that we may infer that the only antecedent difference between fact and foil marks a cause, but in practice there will almost always be many such differences, not all of which will be causally relevant. Moreover, as Mill seems not to have recognized, his own deterministic assumptions entail that there will always be multiple differences as a matter of principle, since any antecedent difference itself marks an effect that must have a still earlier causal difference. (I owe this point to Trevor Hussey.)

In the case of inference, the central problem is to distinguish those differences that are causally relevant from those that are not. In the case of explanation, on the other hand, all the differences that meet the Difference Condition are, by definition, causally relevant. So all of them may be explanatory: the Difference Condition does not entail that there is only one way to explain a contrast. At the same time, however, some causally relevant differences will not be explanatory in a particular context, so while the Difference Condition may be necessary for the causal contrastive explanations of particular events, it is not generally sufficient. For that we need further principles of causal selection.

The considerations that govern selection from among causally relevant differences are numerous and diverse; the best I can do here is to mention what a few of them are. An obvious pragmatic consideration is that someone who asks a contrastive question may already know about some causal differences, in which case a good explanation will have to tell her something new. If she asks why Kate rather than Frank won the prize, she may assume that it was because Kate wrote the better essay, in which case we will have to tell her more about the differences between the essays that made Kate's better. A second consideration is that, when they are available, we usually prefer explanations where the foil would have occurred if the corresponding cause had occurred. Suppose that only Able had a letter from Quine, but even a strong letter from Quine would not have helped Baker much, since his specialities do not fit the department's needs. Suppose also that, had Baker's specialities been appropriate, he would have gotten the job, even without a letter from Quine. In this case, the difference in specialities is a better explanation than the difference in letters. Note, however, that an explanation that does not meet this condition of counterfactual sufficiency for the occurrence of the foil may be perfectly acceptable, if we do not know of a sufficient difference. The explanation of why Jones rather than Smith contracted paresis is an example of this: even if Smith had syphilis in his medical history, he probably would not have

contracted paresis. Moreover, even in cases where a set of known causes does supply a counterfactually sufficient condition, the inquirer may be much more interested in some than in others. The doctor may be particularly interested in causes he can control, the lawyer in causes that are connected with legal liability, and the accused in causes that cannot be held against him.

We also prefer differences where the cause is causally necessary for the fact in the circumstances. Consider a case of overdetermination. Suppose that you ask me why I ordered eggplant rather than beef, when I was in the mood for eggplant and not for beef, and I am a vegetarian. My mood and my convictions are separate causes of my choice, each causally sufficient in the circumstance and neither necessary. In this case, it would be better to give both differences than just one. The Difference Condition could easily be modified to require necessary causes, but I think this would make the condition too strong. One problem would be cases of 'failsafe' overdetermination. Suppose we change the restaurant example so that my vegetarian convictions were not a cause of the particular choice I made: that time, it was simply my mood that was relevant. Nevertheless, even if I had been in the mood for beef, I would have resisted, because of my convictions. In this case, my explanation does not have to include my convictions, even though my mood was not a necessary cause of my choice. (Of course if I knew that you were asking me about my choice because you were planning to invite me to your house for dinner, it would be misleading for me not to mention my convictions, but this goes beyond the conditions for explaining the particular choice I made.) Again, we sometimes do not know whether a cause is necessary for the effect, and in such cases the cause still seems explanatory. But when there are differences that supply a necessary cause, and we know that they do, we prefer them. There are doubtless other pragmatic principles that play a role in determining which differences or combinations of differences yield the best explanation in a particular context. So there is more to contrastive explanation than the Difference Condition describes, but that condition does describe the central mechanism of causal selection.

Since contrastive questions are so common and foils play such an important role in determining explanatory causes, it is natural to wonder whether all why-questions are not at least implicitly contrastive. Often the contrast is so obvious that it is not worth mentioning. If you ask me why I was late for our appointment, the question is why I was late rather than on time, not why I was late rather than not showing up at all. Moreover, in cases where there is no specific contrast, stated or implied, we might construe 'Why P?' as 'Why P rather than not-P?', thus subsuming all causal why-questions under the contrastive analysis. But the Difference Condition seems to misbehave for these

'global' contrasts. It requires that we give a cause of P that finds no corresponding cause in the history of not-Q but, if the foil is simply the negation of the fact, this seems to require that we find a cause of P that finds no corresponding cause of itself, which is impossible, since it is tantamount to the requirement that we find a cause of P that is absent from the history of P.

We can, however, analyse the explanation of P *simpliciter* as the explanation of P rather than not-P. The correct way to construe the Difference Condition as it applies to the limiting case of the contrast, P rather than not-P, is that we must find a difference for events logically or causally incompatible with P, not for a single event, 'not-P'. Thus suppose that we ask why Jones has paresis, with no implied contrast. This would require a difference for foils where he does not have paresis. Saying that he had syphilis differentiates between the fact and the foil of a thoroughly healthy Jones, but this is not enough, since it does not differentiate between the fact and the foil of Jones with syphilis but without paresis. Excluding many incompatible foils will push us towards a sufficient cause of Jones's syphilis, since it is only by giving such a 'full cause' that we can be sure that some bit of it will be missing from the history of all the foils. To explain P rather than not-P, however, we do not need to explain every incompatible contrast. We do not, for example, need to explain why Jones contracted paresis rather than being long dead or never being born. The most we can require is that we exclude all incompatible foils with histories similar to the history of the fact.

One difficulty for this way of avoiding the pathological requirement of finding a cause of P that is absent from the history of P is that there appear to be some facts whose negation also seem to be a single fact (I owe this point to Elliot Sober). Suppose we wish to understand why there are tigers. Here the foil seems simply to be the absence of tigers, and we cannot give a cause of the existence of tigers that is not in the history of tigers. But the existence of tigers is not an event, so this example does not affect my account, which is only meant to apply to the explanation of events. So perhaps the problem does not arise for contrasts whose facts are events and whose foils are either events or sets of events. The Difference Condition will apply to some contrasts that are not explicitly event-contrasts, but not to all of them. Even for P's that are events, however, I am not certain that every apparently non-contrastive question should be analysed in contrastive form, so I am agnostic on the issue of whether all why-questions are contrastive.

Finally, let us compare my analysis of contrastive explanation to the Deductive-Nomological model. First, as we have already noted, a causal model of explanation has the merit of avoiding all the counter-examples to the Deductive-Nomological model where causes are

deduced from effects. It also avoids the unhappy consequence of counting almost every explanation we give as a mere sketch, since one can give a cause of P that meets the Difference Condition for various foils without having the laws and singular premises necesssary to deduce P. Many explanations that the deductive model counts as only very partial explanations of P are in fact reasonably complete explanations of P rather than Q. The excessive demands of the deductive model are particularly striking for cases of compatible contrasts, as least if the deductive-nomologist requires that an explanation of P rather than Q provide an explanation of P and an explanation of not-Q. In this case, the model makes explaining the contrast substantially harder than providing a deductive explanation of P, when in fact it is often substantially easier. Our inability to find a non-contrastive reduction of contrastive questions is a symptom the inability of the Deductive-Nomological model to give an accurate account of this common type of explanation.

There are at least two other conspicuous advantages of a causal contrastive model of explanation over the Deductive-Nomological model. One odd feature of the deductive model is that it entails that an explanation cannot be ruined by adding true premises, so long as the additional premises do not render the law superfluous to the deduction, by entailing the conclusion outright. This consequence follows from the elementary logical point that additional premises can never convert a valid argument into an invalid one. In fact, however, irrelevant additions can spoil an explanation. If I say that Jones rather than Smith contracted paresis because only Jones had syphilis and only Smith was a regular church-goer, I have not simply said more than I need to, I have given an incorrect explanation, since going to church is not a prophylactic. By requiring that explanatory information be causally relevant, the causal model avoids this problem. Another related and unhappy feature of the Deductived-Nomological model is that it entails that explanations are virtually deductively closed: an explanation of P will also be an explanation of any logical consequence of P, so long as the consequence is not directly entailed by the initial conditions alone. (For an example of the slight non-closure in the model, notice that a Deductive-Nomological explanation of P will not also be a Deductive-Nomological explanation of the disjunction of P and one of the initial conditions of the explanation.) In practice, however, explanation seems to involve a much stronger form of non-closure. I might explain why all the men in the restaurant are wearing paisley ties by appealing to the fashion of the times for ties to be paisley, but this might not explain why they are all wearing ties, which is because of a rule of the restaurant. (I owe this example to Tim Williamson.) The contrastive model gives a natural account of this sort of non-closure. When we ask about paisley ties, the

implied foil is other sorts of tie; but when we ask simply about ties, the foil is not wearing ties. The fashion marks a difference in one case, but not in the other.

A defender of the Deductive-Nomological model may respond to some of these points by arguing that, whatever the merits of a contrastive analysis of lay explanation, the deductive model (perhaps with an additional restriction blocking 'explanations' of causes by effects) gives a better account of scientific explanation. For example, it has been claimed that since scientific explanations, unlike ordinary explanations, do not exhibit the interest relativity of foil variation that a contrastive analysis exploits, a contrastive analysis does not apply to scientific explanation (Worrall, 1984, pp. 76–77). It is, however, a mistake to suppose that all scientific explanations even aspire to Deductive-Nomological status. The explanation of why Jones rather than Smith contracted paresis is presumably scientific, but it is not a deduction *manqué*. Moreover, as the example of the thermometer showed, even a full Deductive-Nomological explanation may exhibit interest relativity. I may explain the fact relative to some foils but not relative to others. A typical Deductive-Nomological explanation of the rise of mercury in a thermometer will simply assume that the glass does not break and so while it will explain, for example, why the mercury rose rather than fell, it will not explain why it rose rather than breaking the thermometer. Quite generally, a Deductive-Nomological explanation of a fact will not explain that fact relative to any foils that are themselves logically inconsistent with one of the premises of the explanation. Again, a Newtonian explanation of the Earth's orbit (ignoring the influence of the other planets) will explain why the Earth has its actual orbit rather than some other orbits, but it will not explain why the Earth does not have any of the other orbits that are compatible with Newton's theory. The explanation must assume information about the Earth's position and velocity at some time that will rule out the other Newtonian orbits, but it will not explain why the Earth does not travel in those paths. To explain this would require quite different information about the early history of the Earth. Similarly, an adaptionist explanation for a species's possession of a certain trait may explain why it has that trait rather than various maladaptive traits, but it may not explain why it had that trait rather than other traits that would perform the same functions equally well. To explain why an animal has one trait rather than another functionally equivalent trait requires instead appeal to the evolutionary history of the species, in so far as it can be explained at all.

With rather more justice, a deductive-nomologist might object that scientific explanations do very often essentially involve laws and theories, and that the contrastive model does not seem to account for this. For even if the fact to be explained carries no restricting contrast,

the contrastive model, if it is extended to this case by analysing 'Why P?' as 'Why P rather than not-P?', only requires at most that we cite a condition that is causally sufficient for the fact, not that we actually give any laws. I think, however, that the contrastive model can help to account for the undeniable role of laws in many scientific explanations. To see this, notice that scientists are often and perhaps primarily interested in explaining regularities, rather than particular events (cf. Friedman, 1974, p. 5; though explaining particular events is also important when, for example, scientists test their theories, since observations are of particular events). I think that the Difference Condition applies to many explanations of regularities, but to give a contrastive explanation of a regularity will require citing a law, or at least a generalization, since we here need some general cause (cf. Lewis, 1986, pp. 225–6). To explain, say, why people feel the heat more when the humidity is high, we must find some general causal difference between cases where the humidity is high and cases where it is not, such as the fact that the evaporation of perspiration, upon which our cooling system depends, slows as the humidity rises. So the contrastive model, in an expanded version that applies to general facts as well as to events (a version I do not here provide), should be able to account for the role of laws in scientific explanations as a consequence of the scientific interest in general why-questions. Similarly, although the contrastive model does not require deduction for explanation, it is not mysterious that scientists should often look for explanations that do entail the phenomenon to be explained. This may not have to do with the requirements of explanation *per se*, but rather with the uses to which explanations are put. Scientists often want explanations that can be used for accurate prediction, and this requires deduction. Again, the construction of an explanation is a way to test a theory, and some tests require deduction.

Another way of seeing the compatibility of the scientific emphasis on theory and the contrastive model is by observing that scientists are not just interested in this or that explanation, but in a unified explanatory scheme. Scientists want theories, in part, because they want engines that provide many explanations. The contrastive model does not entail that a theory is necessary for any particular explanation, but a good theory is the best way to provide the many and diverse contrastive explanations that the scientist is after. This also helps to account for the familiar point that scientists are often interested in discovering causal mechanisms. The contrastive model will not require a mechanism to explain why one input into a black box causes one output, but it pushes us to specify more and more of the detailed workings of the box as we try to explain its full behaviour under diverse conditions. So I conclude

that the contrastive model of explanation does not fly in the face of scientific practice.

6. Conclusion

The Difference Condition shows how contrastive questions about particular events help to determine an explanatory cause by a kind of causal triangulation. This contrastive model of causal explanation cannot be the whole story about explanation, since not all explanations are causal explanations or explanations of particular events and since, as we have seen, the choice of foil is not the only factor that affects the appropriate choice of cause. It does, however, give a natural account of much of what is going on in many explanations, and it captures some of the merits of competing accounts while avoiding some of their weaknesses. We have just seen this in some detail for the case of the Deductive-Nomological model. It also applies to the familiarity view. When an event surprises us, a natural foil is the outcome we had expected, and meeting the Difference Condition for this contrast will help to show us why our expectation went wrong. The mechanism of causal triangulation also accounts for the way a change in foil can lead to a change in explanatory cause, since a difference for one foil will not in general be a difference for another. It also shows why explaining 'P rather than Q' is sometimes harder and sometimes easier than explaining P alone. It may be harder, because it requires the absence of a corresponding event in the history of not-Q, and this is something that will not generally follow from the presence of the cause of P. Explaining the contrast may be easier, because the cause of P need not be sufficient for P, so long as it is part of a causal difference between P and not-Q. Again, causal triangulation helps to elucidate the interest relativity of explanation. We express some of our interests through our choice of foils and, by construing the phenomenon to be explained as a contrast rather than the fact alone, interest relativity reduces to the important but unsurprising point that different people are interested in explaining different phenomena. The Difference Condition also shows that different interests do not require incompatible explanations to satisfy them, only different but compatible causes. Moreover, my model of contrastive explanation suggests that our choice of foils is often governed by our *inferential* interests. As I argue extensively elsewhere, the structural similarity between the Difference Condition and the Method of Difference enables us to show why the inductive procedure of 'Inference to the Best Explanation' is a reliable way of discovering causes. Because of this similarity, it can be shown that the hypothesis that would provide the best explanation of our contrastive data is also

the one that is likeliest to have located an actual cause (Lipton, forthcoming). Finally, the mechanism of causal triangulation accounts for the failure of various attempts to reduce contrastive questions to non-contrastive form. None of these bring out the way a foil serves to select a location on the causal history leading up to the fact. Causal triangulation is the central feature of contrastive explanation that non-contrastive paraphrases suppress.

References

Friedman, M. 1974. 'Explanation and Scientific Understanding', *The Journal of Philosophy* **LXXI**: 1–19.

Garfinkel, A. 1981. *Forms of Explanation* (New Haven: Yale University Press).

Hempel, C. 1965. *Aspects of Scientific Explanation* (New York: The Free Press).

Lewis, D. 1986. 'Causal Explanation', in *Philosophical Papers* Vol. II (New York: Oxford University Press), pp. 214–40.

Lipton, P. forthcoming. *Inference to the Best Explanation* (London: Routledge).

Mill, J. S. 1904. *A System of Logic* (London: Longmans, Green and Co).

Ruben, D.-H. 1987. 'Explaining Contrastive Facts', *Analysis* **47.1**: 35–7.

Temple, D. 1988. 'Discussion: The Contrast Theory of Why-Questions', *Philosophy of Science* **55**: 141–51.

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

Worrall, J. 1988. 'An Unreal Image', *The British Journal for the Philosophy of Science*, **35**: 65–80.