Philosophica 51 (1993, 1) pp. 39-54

MAKING A DIFFERENCE

Peter Lipton

1. *Introduction*

An effect is typically explained by citing a cause, but not any cause will do. The oxygen and the spark were both causes of the fire, but normally only the spark explains it. What then distinguishes explanatory from unexplanatory causes? One might attempt to characterise this distinction in terms of intrinsic features of the causes. For example, some causes are changes while others are standing conditions, and one might claim that only the changes explain. Both the spark and the oxygen are causes of the fire, but only the spark is a change, and perhaps this is the reason only the spark explains. On the other hand, one might attempt to characterise the distinction between explanatory and unexplanatory causes in terms of the relation between cause and effect. For example, only some causes are sufficient for their effects, and perhaps only sufficient causes explain.

There is, however, an elementary feature of the distinction between explanatory and unexplanatory causes that neither an intrinsic nor a relational approach are well-suited to capture. This is the so-called 'interest-relativity' of explanation: the very same cause may be explanatory for one person but not for another. When there is a famine in India, an Indian peasant may explain this by citing the drought, while a member of the World Health Organization may instead cite the failure of the Indian government to stock adequate reserves of food (Hart and Honoré, 1985, pp. 35-6).

Why do different people require different explanations of the same effect? A natural thought is that, although they are all asking about the same effect, they are asking different questions about it. This thought can be developed by noticing that many why-questions are contrastive. What

is asked is not simply of the form 'Why this?', but 'Why this *rather than* that?'. The effect is contrasted with a particular foil, and a cause that explains the effect relative to one foil may not explain it relative to another. Thus the drought may explain why there is a famine in India this year rather than in other years, while the failure to build up reserves may explain why there was a famine in India rather than in other countries. So we can usefully investigate the distinction between explanatory and unexplanatory causes by studying contrastive explanation.

This paper will focus on an apparently anomalous feature of contrastive questions. On the one hand, the 'rather than' construction in a contrastive question seems to imply that the contrasted elements — the effect and the foil — are incompatible. Certainly many contrastive questions do have incompatible contrasts. When you ask why Johnson rather than Christie won the race, you know that they could not both have won; when you ask why the mercury in the thermometer rose rather than fell at a certain time, you know that it could not then have done both. Nevertheless, contrasts are often compatible. When someone asks 'Why E rather than F?', he presupposes that E occurred and F did not, but E and F may be independent and so compatible events. A famine in India in one year is compatible both with a famine there in other years and with a famine that year in other countries. Similarly, one may ask why Smith rather than Jones contracted paresis, even though it was obviously not Smith's affliction that protected Jones. Moreover, even when someone asks a contrastive question in the belief that the effect and foil are incompatible, that incompatibility is not presupposed by the question. If I ask you why Jane rather than Frank won the Philosophy Prize, I will not withdraw my question if you tell me that the committee sometimes awards two prizes.

In what follows, I will exploit compatible contrasts, first to criticise two accounts of contrastive explanation and then to motivate a third. I will then account for the apparent tension between the existence of compatible contrasts and the suggestion of incompatibility carried by the 'rather than' construction. This will enable me say something about the point of asking contrastive why-questions. It will help to answer the question, itself contrastive, of why we often ask 'Why E rather than F?' rather than simply 'Why E?'. Finally, I will compare contrastive and Deductive-Nomological explanation.

2. *Three Models*

The incompatibility of effect and foil apparently implied by the 'rather than' construction suggests that contrastive explanation might be analysed as a two-step process (cf. Temple, 1988). First the effect is explained on its own; then it is observed that, since the effect occurred, the foil could not have. This exclusion model fails to account for compatible contrasts. We do not explain why Smith rather than Jones contracted paresis by first explaining why Smith contracted paresis and then showing how this prevented Jones from getting it. But we can explain why Smith rather than Jones contracted paresis, say by pointing out that only Smith had untreated syphilis. The exclusion model also fails to give an adequate account of incompatible contrasts, since it cannot show how a change of foil may make a previously explanatory cause unexplanatory, or *vice versa*. A explanation of why Johnson rather than Christie won the race may not explain why Johnson rather than Williams won. If the exclusion model were correct, however, anything that explains an effect relative to one incompatible foil would also explain it relative to another.

David Lewis (1986, pp. 229-30) has given an alternative account of contrastive explanation. According to him, to explain why E rather than F we must cite a cause of E that would not have been a cause of F, had F occurred. In his example, we can explain why Lewis went to Monash rather than to Oxford in 1979 by pointing out that Monash invited him, since the invitation to Monash was a cause of his going there but of course would not have been a cause of his going to Oxford, had he done so. On the other hand, while his desire to go to a place where he has good friends was also a cause of going to Monash, it does not explain why he went there rather than to Oxford, since he also has good friends there. His desire thus would have been a cause of going to Oxford, had he gone there instead.

Lewis's model is an improvement on the exclusion model in several respects. In particular, it leaves room for the fact that a cause that explains an effect relative to one foil may not explain it relative to another. The fact that the race was held at high altitude may explain why Johnson rather than Christie won, if only Johnson does particularly well under such conditions, but will not explain why Johnson rather than Williams won, if Williams is also a high-altitude specialist. To the credit of Lewis's model, under these suppositions the high altitude would not have

been a cause of Christie winning, but would have been a cause of Williams winning. Moreover, as we will eventually see, the counterfactual element that Lewis invokes does bring out an important aspect of contrastive explanation.

Nevertheless, like the exclusion model, Lewis's account falls to compatible contrasts. While the exclusion model makes the explanation of compatible contrasts impossible, Lewis's model makes them too easy. For where effect and foil are compatible, few of the causes of the effect would have been causes of the foil. We cannot explain why Smith rather than Jones contracted paresis by observing that Smith had syphilis, if Jones did as well. Yet Smith's syphilis is a cause of his paresis that would not have been a cause of Jones's paresis, and so satisfies Lewis's requirements. Similarly, the failure to build up reserves of food in India will not explain why there was a famine in India rather than in Egypt, if Egypt didn't have reserves either, yet India's failure to build up reserves would not have been a cause of a famine in Egypt, had there been one.

This difficulty for Lewis's account extends to incompatible contrasts. For suppose that Lewis could not have gone both to Monash and to Oxford in 1979, but that he received an invitation from both. In this case, the invitation to Monash clearly would not explain why he went there rather than to Oxford, but the invitation to Monash still satisfies Lewis's conditions. It was a cause of going to Monash and would not have been a cause of going to Oxford. It is also worth noting that the account can not be saved by switching from causal token to causal type. With this modification, the account would be that an explanation of E rather than F is a cause of E of a type such that there would have been no token of that type causing F, had F occurred. This would correctly rule out causes such as the invitation to Monash where Oxford invited as well, at least if we assume that Lewis only goes where he is invited. In this case, an invitation would have been a cause of his going to Oxford. But the modified account is now too restrictive, ruling out the perfectly good explanation in terms of the invitation from Monash in the case where only Monash invites. For supposing again that Lewis only goes where he is invited, even if in actuality only Monash invited, an invitation still would have been a cause of his going to Oxford.

I want now to sketch a third account of contrastive explanation that improves on both the exclusion and Lewis models. If we focus on compatible contrasts, there is a striking analogy between contrastive expla-

nation and Mill's method of difference, his version of the controlled experiment. According to Mill, one of the most common and powerful ways of inferring causes from effects exploits contrasting instances, where the effect occurs in one but not in the other. If the two instances share every possible cause except one, and this one occurs only in the instance where the effect also occurs, we may infer that this circumstance is a cause. If we wish to find a cause of paresis, and the only plausibly relevant difference between the medical histories of Smith and Jones is that only Smith had syphilis, then we are entitled to infer that syphilis is a cause of paresis.

The role of the method of difference is ostensibly different from that of contrastive explanation. In one we infer causes from effects; in the other we explain effects by causes. There is nevertheless a strong structural similarity between the two activities. In both cases, we begin with an effect and a foil, and in both cases we go on to look for a prior difference. This suggests that explaining a contrast requires a cause that *made the difference* between effect and foil. In many cases, what this amounts to is finding a cause in the instance where the effect occurs where there is no cause of the same type — no corresponding token — in the foil instance.

This difference model gives the right answers for many compatible contrasts. If only Smith has syphilis, this explains why he rather than Jones contracted paresis, since Smith's syphilis is a cause of his paresis and there is no corresponding token (Jones's syphilis) in the case of Jones. The drought in India explains why there was famine there that year rather than in other years, since there was no drought in those years, but it does not explain why there was famine in India that year rather than in other countries which also suffered a drought. The difference model also works for many incompatible contrasts (as does Mill's method of difference). The invitation from Monash explains why Lewis went there rather than to Oxford just in case he did not also receive an invitation from Oxford. If Johnson took steroids, this explains why he rather than Christie won the race just in case Christie did not enjoy the same illegal benefit. Finally, the difference model clearly captures the fact that a cause that explains an effect relative to one foil may not do so relative to another. Johnson's steroids explains why he won rather than someone who did not take steroids, but not why he won rather than someone else who did.

The difference model differs from the exclusion model in not requiring that the effect preclude the foil, and it differs from Lewis's model in emphasising an actual rather than a counterfactual difference (though as we will see below, there is also a counterfactual element in contrastive explanation). It provides a better picture of what is going on in many contrastive explanations, but it is also clearly too simple as it stands. There are a number of questions a fuller account ought to answer. First, what restrictions are there on the type under which a cause can fall for the purposes of contrastive explanation? Second, how is the notion of 'corresponding token' to be defined? The difference model requires that we have both the presence of one token (the cause) and the absence of another, but obviously not any absence will do. The fact that Cambridge did not invite Lewis is obviously not enough to make the invitation to Monash explain why he went to Monash rather than to Oxford. In practice it is usually clear which token is relevant, but one would like a principled analysis.

A third question arises because of cases where, although the explanatory cause did intuitively 'make the difference' between effect and foil, the notion of corresponding token does not seem applicable. The rise in temperature made the difference between the rising and the falling of the mercury in the thermometer and so explains this contrast (though not why the mercury rose rather than breaking the glass), but this is not a case where there was rising temperature in one actual instance but not in another, so the notion of a corresponding token gets no purchase. Another type of case where the corresponding token analysis seems inapplicable is one where the cause that makes the difference is one that produces a perturbation and where the foil is what would have occurred without this interference. (I owe this point to Jonathan Vogel.) The particle was deflected rather than moving in a straight line because it passed through a particular field: this field made the difference, but not because there was a field in one instance but not in another. The difficulty in cases such as these is not that effect and foil are incompatible since, as we have seen, the model handles many such cases in a straightforward way. The difficulty seems rather to be that, in the troublesome cases, we do not have two distinct and actual instances, one in which the phenomenon occurs and one in which it does not. As a consequence, it is difficult to see how we can here speak of a corresponding token that is absent but might have been present. All contrasts have non-actual foils, but the

difference model (like Mill's method) only works smoothly where we also have two actual instances. The third question is thus how the notion of making a difference is to be analysed when this condition is not satisfied.

A fourth question is whether, in the typical cases where it does make sense to speak of the absent token, mere absence is enough, or whether the account ought to impose some further requirement, such as that the corresponding token would have been a cause of the foil, had the foil occurred. The final question I will mention concerns multiple differences. In real life there will seldom if ever be only a single causal difference between fact and foil, so the model needs to say what explanation requires when there are several, both where these differences work in the same direction and when they work against each other.

These are difficult questions, though their difficulty does not make either the exclusion or Lewis models any more attractive. Rather than attempt partial answers here, however, I want now to consider another perspective on contrastive explanations, which yields a different but complementary picture of their function. Like what has come before, this picture will be motivated by the tension between compatibility and incompatibility in the contrasts we query. I have so far emphasised the existence of compatible contrasts, both to criticise various accounts of contrastive explanation and to motivate one of my own. But we still have the suggestion of incompatibility that the 'rather than' construction carries, even in cases where the explicit contrasts are compatible. By seeing why this is so, we will be able to say something further about the mechanism and point of contrastive explanations.

3. *Latent Incompatibility and Backward Counterfactuals*

The reason contrastive questions carry the implication of incompatibility even when the contrasts are compatible is that the actual foil is typically a surrogate for a counterfactual claim about the effect. To make this clearer, let us return for a moment to Mill's method of difference. Mill tells us that, if we want to find a cause of an effect, we should look among the differences between an instance where the effect occurs and a similar instance where it does not. This is good advice, but there is a sense in which the experiment is a surrogate for the one we would really like to perform, if only we could. That ideal experiment would be not a

comparison of one instance with another, but a comparison of an instance with itself! That is, the experiment we would really like to perform would start with the instance in the actual world where the effect occurs and compare it with the very same instance in nearby possible worlds where the effect does not occur. We would like to see what would have been different about Smith, if *he* had not had paresis. Of course this is something we cannot see, our powers of perception being limited to small portions of the actual world. So we do the next best thing: we find or construct a *doppelganger* instance in the actual world to serve as a surrogate for the possible world, and use it to assess the counterfactual. We thus establish a connection between what we can observe and the counterfactual that interests us.

I think the situation is similar in the case of contrastive explanation. When we ask questions such as why Smith rather than Jones contracted paresis, our underlying interest often really concerns a contrast about Smith alone. That is, we are really asking a certain type of question about why Smith had paresis rather than what would have been the case, had *Smith* not had paresis, an obviously incompatible contrast. The talk about Jones is a way of getting at a certain type of question about Smith. Thus we see why a contrastive question retains the feeling of incompatibility even when the explicit contrast is compatible. That contrast is compatible, but it is also a surrogate for an underlying and incompatible contrast between what was and what might have been. This is the reason for the sense of incompatibility that the 'rather than' construction carries even when effect and foil are compatible. It also helps to explain why it is so often supposed that contrasts must be incompatible, in spite of the obvious counterexamples.

This resolution does, however, raise a further question. If the compatible foil is a surrogate for the incompatible contrast that really interests us, why do we take this detour? If what I have said so far is along the right lines, a contrastive question, whether or not the contrast is compatible, is really a certain sort of question about what made the difference between the effect's occurring and its not occurring. Why then do we bother asking contrastive questions with specific foils at all? Why don't we stick to the global form 'Why E rather than not-E?' or, to save ourselves some breath, why not simply 'Why E?'? In other words, what is the point of asking contrastive questions?

In the case of Mill's method, the reason for the actual and specific

foil is clear: in the context of experiment, we need something observable and hence actual. I have argued elsewhere (Lipton, 1991, ch. 5) that this motivation carries over into some of the uses of contrastive explanation, since we often infer causes by means of an inference to the best explanation of contrasts in the evidence. Potential explanations of why an effect occurred in one instance but not in another are a guide to the likeliest cause. This cannot, however, be the whole story, since explanation is not only a tool for inference and the contrasts we ask about are not restricted by the limits of observability. So we need to say something more about the point of contrastive explanation.

Consider the global contrastive question, 'Why E rather than not-E?'. Here we are asking what made the difference between E and not-E, which involves asking a counterfactual question about how things would have been, had E not occurred. This is, however a peculiar and awkward question, because the conditional is a 'backward' counterfactual. It asks how things would have different earlier, if something had been different later. This is unlike typical counterfactuals, which are forward-directed, following the direction of causation. We consider how things would have been later, if they had been different earlier. For example, we say that if a certain cause hadn't occurred, a certain effect wouldn't have occurred either. Now while all counterfactuals suffer from a vagueness that needs to be settled by the context of use, many of these ordinary 'forward' counterfactuals are clearly true in the context in which they appear. It may be uncontroversially true that, for example, if Johnson had never trained, he wouldn't have won the race, or that if the match hadn't been struck, it would not have lit. But as David Lewis (1979, pp. 32-5) has observed, backward counterfactuals appear to suffer from far worse vagueness or indeterminacy. If Johnson had not won the race, who knows what earlier things would have been different? If the match had not lit, who knows whether it still would have been struck?

I think Lewis is right to say that backward counterfactuals tend to suffer greater indeterminacy than forward ones. But this difference leads him to take a very strict line on backward counterfactuals. Under what he calls the 'standard resolution' of vagueness, he claims that a backward counterfactual is true if and only if its consequent is itself true (Lewis, 1979, p. 35). Thus he holds that, if things had been different today, they would have nevertheless remained unchanged in the past.

Lewis has overreacted. I will here make just two brief objections to

his view that things would have been the same earlier even if they had been different later. The first is that this view has a very implausible consequence; the second is that it is in tension with Lewis's own position on contrastive explanation. (For an extended critique of Lewis's treatment of backward counterfactuals, see Bennett, 1984.) The implausible consequence is that it makes the laws of nature incredibly fragile. Lewis's position is developed under the assumption that the actual world is deterministic. Given this assumption, saying that the past would remain the same even if the present had been different requires what Lewis calls a 'miracle'. If the present had been different in any respect, the laws of the actual world would been violated. The point is not that the laws would have exceptions, but that there would be different laws. So, according to Lewis, if you had not seen this article, the laws of nature would have been different. While he embraces this consequence, I find it sufficiently incredible to refute his position.

The second objection is that Lewis himself needs backward change for his account of contrastive explanation. (I owe this point to Philip Bricker and Jonathan Vogel.) Recall his account: to explain why E rather than F, we must cite a cause of E that would not have been a cause of F, had F occurred. Thus, unlike the invitation from Monash, Lewis's desire to be with friends does not explain why he went to Monash rather than to Oxford, since the desire would have been a cause of his going to Oxford, had he gone there instead. That is, according to Lewis, if he had gone to Oxford, the desire would have been a cause. But this is a backward counterfactual, and one that implies just the sorts of change Lewis's own account of such counterfactuals proscribes. Had Lewis gone to Oxford, caused in part by desire to be with friends, things would not have been just the same up to a last-minute miracle. He would have written different letters, spoken with different people, booked different tickets, and so on. The problem for Lewis here is not simply that the miracle scenario is incredible, supposing as it does that he would have acted just as if he were going to Monash, but somehow found himself in Oxford. For even if we grant the unchanged past, Lewis loses the discrimination his counterfactuals are supposed to provide. If things would have been just the same, up to a last-minute miracle, had Lewis gone to Oxford, then who is to say that the desire would have been a cause while the invitation would not have been?

So I take it that Lewis is wrong about backward counterfactuals. The

standard resolution is often not one that holds the past constant at the cost of the laws of nature; rather it seems to hold the laws constant to consider how the past would have been different. One such context is an explanatory one, when we consider how things would have been different, had E not occurred. But this leaves us with the problem with which Lewis began: the extreme indeterminacy of these backward counterfactuals. What I will now suggest is that reducing this indeterminacy is one of the points of making our why-questions contrastive.

4. *Focussing and Fixing*

Specific contrasts help with the problem of the indeterminacy of the question of how things would have been different earlier, had E not occurred, in at least two ways. The first is by 'focussing' the antecedent; the second is by 'fixing' part of the past. The effects we want explained are usually relatively specific, so the global foil — not-E — is usually relatively general. There are many ways in which something can fail to happen, and this is one reason we are sometimes at a loss to evaluate the relevant counterfactuals. There are so many ways Johnson might have lost the race that we don't know how things would have been different if he had lost. Think, for example, of all the differences in the other runners that might have resulted in him losing. Similarly, there are so many things Lewis might have done instead of going to Monash that we don't know how things would have been different if he hadn't gone there. Specific contrasts help here, by providing what Alan Garfinkel (1981, p. 30) has called a 'limited negation'. We are in a better position to say how things would have been different had Johnson lost, if we stipulate that he lost to Christie, and we are in a better position to say how things would have been different if Lewis had not gone to Monash if we stipulate that he went instead to Oxford.

The second way specific contrasts help to resolve indeterminacy is by fixing parts of the causal history of the effect, namely those parts that find an echo in the foil-instance. Had there not been a famine in India, perhaps it would have been because there had been no drought, or perhaps because the government would have had sufficient food reserves. This indeterminacy is eliminated if we ask why there was a famine there this year rather than in other years, or if we ask why there was a famine

that year in India rather than in other countries that also suffered drought. In the first case we fix the absence of reserves; in the second the drought. By holding part of the past fixed, we help to resolve the vagueness about what would have to have been different. Similarly, when we ask why Smith contracted paresis, we want to know how the past would have been different, had Smith not had paresis. By using Jones as a foil, we reduce the vagueness of the question. It is as if we asked the following: If Smith had not had paresis, but Smith's past were held fixed wherever it is similar to Jones's past, how would Smith's past have differed from the way it actually was? By choosing a sensible contrast, we fix enough to give the relevant backward counterfactuals a truth-value. One of the neat features of this technique is that it enables us to fix with our question more of the past than we initially know. Every similarity between effect and foil is held fixed, even though we do not know what all the similarities are.

The link between contrastive questions and backward counterfactuals suggests that the difference model might be augmented with a counterfactual component. Although I cannot develop this idea here, it seems promising. It might, for example, help us to extend the difference model to cover the difficult cases of the mercury and the deflected particle. Such a development would bring us closer to Lewis's own account of contrastive explanation, but it would enable us to avoid the difficulties we found in that account. One of those was illustrated by the case where both Monash and Oxford invited. As we saw, Lewis's account incorrectly allows the invitation from Monash to explain why he went there rather than to Oxford in that case. The counterfactuals I have been considering do not have this consequence, since when we consider what would have been the case earlier had Lewis gone to Oxford, we hold fixed the invitation from Monash, since there was also an invitation from Oxford. Again, whereas Lewis's account fails for compatible contrasts, this difficulty can be avoided by treating the compatible contrast as a surrogate for one that is incompatible in the way I have sketched.

If one of the benefits of asking contrastive questions is that it reduces the vagueness or indeterminacy of the salient counterfactuals, this casts some further light on the use of compatible foils. I have suggested that we can resolve the apparent tension between the use of compatible foils and the implication of incompatibility carried by the 'rather than' locution by seeing compatible foils as surrogates for an incompatible contrast

between the effect and its absence. That is, the question, 'Why E rather than F?' is a way of asking a certain type of question about what made the difference between E and not-E. But this resolution raised the further question of why we should then so often make this detour through a compatible foil. The role of foils in resolving the vagueness of the backward counterfactual now suggests an answer. By using a compatible foil, we are sometimes able to give our counterfactual question greater determinacy than an incompatible foil would allow, and we are able to pick which parts of the history of the effect we wish to hold fixed in a way that meets our explanatory interests. When we ask why there was a famine in India this year, we want to know how things would have been different, had there been no famine. As it stands, however, this question is hopelessly indeterminate, and there seems here to be no incompatible contrasts that would provide the focussing and fixing we need. The job is neatly done, however, by a compatible contrast, whether it be the absence of famine in India in other years or the absence of famine in other countries this year.

The need for determinacy also provides one answer to the more general question about the point of making why-questions contrastive. When we ask why something happened, we are often asking a question about how things would have been different earlier, had the effect not occurred. Because of the vagueness of backward counterfactuals, however, this bare question may not be sufficiently well-posed to permit a sensible answer. By making our question contrastive, we can often resolve the vagueness sufficiently to avoid this problem and in a way that meets our explanatory interests. This explains why contrastive questions can often do a job that the simple question 'Why E?' cannot. The way in which foils resolve vagueness also helps to explain why the difference model should give a good picture of the way many contrastive explanations work. By looking for a cause of the effect that is of a type absent in the case of the foil, we are generally led to something that made the difference between the effect happening and not, subject to the focussing and fixing that the foil provides.

5. *Contrastive and Deductive-Nomological Explanation*

I will end this essay in the traditional way, by making some comparisons

with the Deductive-Nomological model. Because the account of contrastive explanations I have discussed is a causal model of explanation, it inherits the general advantages such models have over a deductive account. For example, they block the unexplanatory deductions of causes from effects. It must however be said that the Deductive-Nomological model also seems to have several advantages over most causal models. One of the most striking of these is that a Deductive-Nomological explanation provides a sufficient condition for the effect that is explained. This is an advantage because deductive sufficiency can be one explanatory goal: it satisfies our desire to show that the phenomenon that concerns us in some sense had to occur as it did. The fact that Deductive-Nomological explanations require such sufficiency can also, however, be seen as a disadvantage, because it is such a difficult requirement to meet. Once we descend from ideal systems to the complexity of real-world events and processes, it is in general enormously difficult to provide a Deductive-Nomological argument that is both explanatory and contains only true premises. Just try it for, say, Lewis's trip to Monash.

This brings out one of the central advantages of contrastive explanations, for they enable us to provide something like 'sufficiency on the cheap'. By saying what made the difference between the effect happening and not, a contrastive explanation gives us a kind of sufficient condition. This kind of sufficiency is much easier to achieve than the kind the Deductive-Nomological model requires, because of the mass of shared material that the backward counterfactual holds fixed, in the way I have described. Contrastive explanations thus enable us to satisfy (or nearly satisfy) the sufficiency ideal without setting standards we cannot in practice meet.

There are additional advantages to contrastive explanation that apply in cases where a Deductive-Nomological explanation is also available: contrastive explanation is not merely a second-best option, forced upon us by the difficulties in meeting the Deductive-Nomological conditions. Notice first that, even when we have a Deductive-Nomological explanation of E, we may not have an explanation of E rather than F. Consider a typical Deductive-Nomological explanation of the mercury's rise in a thermometer. This would explain why the mercury rose rather than falling, but it would probably not explain why it rose rather than breaking the glass. For the explanation would simply assume that the glass did not break, rather than explaining this. Because of the great difficulty in

producing deductively sufficient conditions, most Deductive-Nomological explanations must include singular premises that are not antecedent conditions, but rather stipulate that certain things remain fixed at the time of the effect, thus suppressing the cause that the contrastive question demands. A Deductive-Nomological argument that cites a cause of the effect may nevertheless not cite a cause that explains the contrast at issue.

The final advantage of contrastive explanation that I will mention applies even in those cases where the Deductive-Nomological explanation does include the explanatory cause, and it has to do with the purposes for which we request explanations. An apparent advantage of explanations that provide deductively sufficient conditions is that they seem automatically to provide something that made the difference between E and not-E. As such, they might appear to avoid the usual difficulties faced by backward counterfactuals and so also to avoid the need for the focussing and fixing that contrastive questions supply. If the explanation entails E, then we know that the explanation would not have obtained, had E not occurred. This advantage, however, is only apparent. The problem is familiar: if the conclusion of a deductively valid argument were false, we know that at least one of the premises would be false as well, but we don't know which. Thus, in the case of a Deductive-Nomological argument, Lewis would say that a lawlike premise would fail, while I would give up one of the singular premises. And even if we agree to hold the laws constant, a Deductive-Nomological explanation, unlike a good contrastive explanation, does not tell us which singular premise would fail. Suppose that my car is belching thick, black smoke. Wishing to correct the situation, I naturally ask why it is happening. Now imagine that God (or perhaps an evil genius) presents me with a full Deductive-Nomological explanation of the smoke. This may not be much help. The problem is that many of the causes of the smoke are also causes of the car's normal operation. Were I to eliminate one of these, I might only succeed in making the engine inoperable. By contrast, an explanation of why the car is smoking *rather than running normally* is far more likely to meet my diagnostic needs. Of course diagnosis and repair is only one of the motives for asking why-questions, but an investigation of the others would reveal still further reasons why we so often make our questions contrastive.

Cambridge University

ACKNOWLEDGEMENTS

I am grateful for the comments of Chris Daly, Mike Dixon, Michael Gaylard and Jonathan Vogel.

REFERENCES

Bennett, J. 1984. "Counterfactuals and Temporal Direction." *The Philosophical Review* XCIII, No. 1, pp. 57-91.

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

Hart, H. L. A. & Honoré, T. 1985. *Causation in the Law* second edition. Oxford: Oxford University Press.

Lewis, D. 1979. "Counterfactual Dependence and Time's Arrow." *Nous* 13, reprinted in Lewis (1986) pp. 32-52.

Lewis, D. 1986. *Philosophical Papers* Volume II. New York: Oxford University Press.

Lipton, P. 1991. *Inference to the Best Explanation*. London: Routledge.

Temple, D. 1988. "The Contrast Theory of Why-Questions." *Philosophy of Science* 55, No. 1, pp. 141-151.