Chapter 3

The causal model

Fact and foil

According to the causal model of explanation, to explain a phenomenon is simply to give information about its causal history (Lewis 1986) or, where the phenomenon is itself a causal regularity, to explain it is to give information about the mechanism linking cause and effect. If we explain why smoking causes cancer, we do not give a cause of this causal connection, but we do give information about the causal mechanism that makes it. Not only is the causal model of explanation natural and plausible, but it avoids many of the problems that beset the other views we have canvassed. It provides a clear distinction between understanding why a phenomenon occurs and merely knowing that it does, since it is possible to know that a phenomenon occurs without knowing what caused it. Moreover, the model draws this distinction in a way that makes understanding unmysterious and objective. Understanding is not some sort of super-knowledge, but simply more knowledge: knowledge of causes.

Unlike the unification and necessity models of explanation, the causal model makes it clear how something can explain without itself being explained or already understood, and so has no difficulty accounting for the possibility of explanation in the face of the regress of whys. One can know a phenomenon's cause without knowing the cause of that cause. And unlike the reason model, which requires that an explanation provide a reason to believe the phenomenon occurs, the causal model accounts for the legitimacy of self-evidencing explanations, where the phenomenon being explained is also an essential part of the evidence for the explanation. The causal model also avoids the most serious objection to the familiarity model, since a phenomenon can be common and unsurprising, even though we do not know its cause. Finally, it avoids many of the objections to the deductive-nomological model. Ordinary explanations do not have to meet the requirements of that model, because one need not give a law to give a cause, and one need not know a law to have good reason to believe that a cause is a cause. As for the over-permissiveness of the deductive-

nomological model, the reason recession explains red-shift but not conversely is that causes explain effects and not conversely; the reason a conjunction does not explain its conjuncts is that conjunctions do not cause their conjuncts; and the reason the sun explains the warmth, while not being warmed does not explain not being in the sun, is that the sun causes an object to warm, but not being warmed does not cause something to be in the shade.

There are three natural objections to the causal model of explanation. The first is that we do not have a fully adequate analysis of causation, and not through want of trying (cf. Sosa and Tooley 1993). This, however, is no reason to abjure the model. The notion of causation is indispensable to philosophy, ordinary life and much of science, we know a good deal about the relation without a full philosophical account, and if we wait for a fully adequate analysis of causation before we use it to analyze other things we will probably wait forever. I will not, in this book, say anything on the large topic of the nature of causation, but trust that what I do say about the role of causation in explanation and inference holds for the causal relation as it is. Applying the moral of the why-regress reflexively, we do not need fully to explain causation in order to use causation to explain other things, in this case, the nature of explanation itself.

The second objection to the causal model of explanation is simply that there are non-causal explanations. Mathematicians and philosophers, for example, give explanations, but mathematical explanations are never causal, and philosophical explanations seldom are. A mathematician may explain why Godel'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. (Some philosophical explanations are, however, broadly causal, such as the explanations of inferential and explanatory practices that we are considering in this book.) There are even physical explanations that seem non-causal. I am particularly fond of two examples. First, suppose that a bunch of 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 moment during the sticks' descent. Why are appreciably more of them near the horizontal axis than near the vertical, rather than in more or less equal numbers near each orientation as one might 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 indefinitely many horizontal orientations, but only two vertical orientations. Or think of the shell that the ends of that stick trace as it takes every possible orientation. The areas that correspond to near the vertical are caps centered on the north and south poles formed when the stick is forty-five degrees or less off the vertical, and this area is substantially less than half the surface area of the entire sphere. Another way of putting it is that the explanation why more sticks are near the horizontal than near the vertical 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.

My second example of a lovely non-causal explanation concerns reward and punishment, and is based on the influential work in cognitive psychology by Daniel Kahneman and Amos Tversky (Kahneman *et al*. 1982: 66–8), work we will return to in chapters 7 and 8. Flight instructors in the Israeli air force had a policy of strongly praising trainee pilots after an unusually good performance and strongly criticizing them after an unusually weak performance. What they found is that trainees tended to improve after a poor performance and criticism; but they actually tended to do worse after good performance and praise. What explains this pattern? Perhaps it is that criticism is much more effective than praise. That would be a causal explanation. But this pattern is also what one should expect if neither praise nor criticism had any effect. It may just be regression to the mean: extreme performances tend to be followed by less extreme performances. If this is what is going on, we can have a lovely explanation of the observed pattern by appeal to chance (or the absence of causal influence) rather than any cause. (This example ought to give pause to parents who are too quick to infer that punishing children for bad behavior is more effective than rewarding them for good behavior.)

The existence of non-causal explanations show that a causal model of explanation cannot be complete. One reaction to this would be to attempt to expand the notion of causation to some broader notion of 'determination' that would encompass the non-causal cases (Ruben 1990: 230–3). This approach has merit, but it will be difficult to come up with such a notion that we understand even as well as causation, without falling into the relation of deductive determination, which will expose the model to many of the objections to the deductive-nomological model. For the time being at least, I believe that the causal view is still our best bet, because of the backward state of alternate views of explanation, and the overwhelming preponderance of causal explanations among all explanations. Nor does it seem ad hoc to limit our attention to causal explanations. The causal view does not simply pick out a feature that certain explanations happen to have: causal explanations are explanatory *because* they are causal.

The third objection is that the causal model is too weak or permissive, that it underdetermines our explanatory practices. Let us focus on the causal explanation of particular events. We may explain an event by giving some information about its causal history, 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. In a particular context, most information about the causal history of a phenomenon is explanatorily irrelevant, so explaining cannot simply be giving such information. This is an important

objection, but I prefer to see it as a challenge. How can the causal model be developed to account for the causal selectivity of our explanatory practices? The rest of this chapter is a partial answer to this question. The answer is interesting in its own right, and it will also turn out to be a crucial tool for developing and assessing an account of Inference to the Best Explanation, the central project of this book.

What makes one piece of information about the causal history of an event explanatory and another not? The short answer is that the causes that explain depend 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 natural way to show how interests help us to select from among causes is to reveal additional structure in the phenomenon to be explained, structure that varies with interest and that points to particular causes. The idea here is that we can account for the specificity of explanatory answer by revealing the specificity in the explanatory question, where a difference in interest is an interest in explaining different things. 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 Carl Hempel has observed, however, we do not explain events, only aspects of events (1965: 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. Which aspect we ask about depends on our interests, and 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 the interest relativity of explanation can be accounted for 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: 28–41; van Fraassen 1980: 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 *simpliciter*, but only for example why they turn yellow in November rather than in January, or why they turn yellow in November rather than turn blue.

The contrastive analysis of explanation is extremely natural. We often pose our why-questions explicitly in 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, then, 3-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. 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, an explanation of why Able rather than Baker got the philosophy job may not explain why Able rather than

Charles got the job, and an explanation of why the mercury in a thermometer rose rather than fell may not explain why it rose rather than breaking the glass. 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 often 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.

The role of contrasts in explanation will not account for all the factors that determine which cause is explanatory. For one thing, I do not assume that all why-questions are contrastive. For another, even in the cases of contrastive questions, the choice of foil is not, as we will see, the only relevant factor. Nevertheless, it does provide a central mechanism, so I want to try 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: 40; Temple 1988: 144; Ruben 1987). 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 (Scriven 1959: 480; Hempel 1965: 369–70; van Fraassen 1980: 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 fact and foil would yield a sensible why-question but, as we will

see, it is not necessary to restrict contrastive why-questions to incompatible contrasts to distinguish sensible questions from silly ones.

The existence of compatible contrasts is somewhat surprising, since the 'rather than' construction certainly appears to suggest some sort of incompatibility (Carroll 1997, 1999). I think there are a number of reasons why we find the 'rather than' construction natural even when the P and Q in 'Why P rather than Q?' are compatible. As I have already mentioned, we often ask a contrastive question when cases we expected to turn out the same in fact turn out differently. In this case, though fact and foil are compatible, we are also interested in the contrast between turning out the same and turning out differently, and *this* contrast is obviously incompatible. In other cases, when we ask 'Why P rather than Q?' we may also be interested in why things turned out one way rather than the other way around – the contrast between (P & not-Q) and (Q & not-P) – which is again an incompatible contrast, even though P and Q are compatible (Jones 1994). A third underlying incompatible contrast in cases of compatible fact and foil is between the fact and its negation: the foil Q is a device for asking a certain kind of question about the contrast between P and not-P. In ways we will investigate, the choice of foil serves to focus on which aspect of the fact we are interested in explaining, in effect by specifying which way of the fact not occurring is of interest. The foil provides, in Alan Garfinkel's words, a 'limited negation' (1981: 30).

So we see there are a number of reasons why the suggestion of incompatibility carried by 'rather than' is apt even when fact and foil are compatible, since there may be incompatible contrasts at play beneath the surface. Now in such cases one could take the terms of those underlying contrasts to be the 'real' facts and foils, so making the questions about an incompatible contrast after all, but I prefer to hold on to P and Q in the surface structure of the question as fact and foil, and these will often be compatible. This way of proceeding gives us a univocal structure for contrastive questions and will make my subsequent analysis more perspicuous, but this is not to deny that the underlying contrasts are real as well.

So that is my first observation: fact and foil may be compatible. 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 or decent explanation of a fact alone, we can see that the requirements for explaining a fact are different from the requirements for explaining a contrast. Of course intuitions about particular non-contrastive explanations may vary, so that for example what to one

person seems no explanation at all will to another seem a genuine though very weak explanation. My hope is that this will not matter in what follows. A good explanation of P rather than Q that is no explanation at all of P alone is a particularly dramatic difference, but for my purposes it is probably enough if you find that the one explanation is considerably better than the other, even if in my presentation I speak in starker terms.

My second observation, then, is that explaining a contrast is sometimes easier than explaining the fact alone, in the sense that an explanation of 'P rather than Q' is not always an explanation of P (cf. Garfinkel 1981: 30). This is particularly clear in examples of compatible contrasts. Jones's syphilis does not, it seems to me, 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. (For a different view, see Carroll 1999.) 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 a very good answer to the simple question, 'Why did you order eggplant?' (even if one does not hold that only sufficient causes explain non-contrastive facts (cf. Hitchcock 1999: 603–4)). 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 to 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: we 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, since it may not include information about the foil.

Failed reductions and false differences

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 both easier and 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'. More interestingly, a deductive-nomological explanation of P is not always sufficient to explain 'P rather than Q', for any incompatible 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; Carroll 1997, 1999). 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. To do so, the conjunctive view should be taken to entail that explaining a conjunction at least requires explaining each conjunct; that 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.

The conjunctive view 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. If your horse is lame and mine isn't, that explains why my horse won rather than yours, but it does not explain why my horse won, there being many other fit horses in the race. The conjunctive view makes contrastive explanation too hard. Somewhat surprisingly, it also makes it too easy, 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. (The deductive-nomological model is nearly but not entirely closed, since it requires that the law premise be essential to the deduction and this condition will not be satisfied by every 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, and that an explanation of P relative to one contrast is not in general an explanation of P relative to another. The conjunction in these cases 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: 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 analyze '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 color 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 fall 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.

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. Intuitively, it seems that to explain a contrast requires citing a cause of the fact that marks a difference between fact and foil. But how is this difference to be analyzed? In the remainder of this section we consider two approaches that do not seem quite right; in the next section I shall try to do better.

David Lewis has given an interesting account of contrastive explanation that does not depend on paraphrasing the contrastive question and that does give one sense of a cause marking a difference between fact and foil. 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: 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 counterfactual 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, since Jones's syphilis would not have been a cause of Smith's paresis, had Smith contracted paresis, yet in this case Jones's syphilis would not explain why he 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.

A second plausible attempt to say what contrastive explanation requires in terms of a cause that marks a difference between fact and foil appeals to a probabilistic notion of favoring. Such an account could take various forms, but a simple version would say that an explanation of why P rather than Q must cite a cause of P that raises the probability of P without raising the probability of Q, where the probabilities are construed as physical chances (van Fraassen 1980: 146–51; Hitchcock 1999: 597–608). (Construing the probabilities instead as degrees of belief is another option, but we then face something analogous to the old evidence problem for Bayesian accounts of confirmation, since P is typically already known when the why-question is

posed.) This favoring criterion rightly counts Lewis's invitation to Monash as an explanation of why he went there rather than to Oxford in a situation where only Monash invites, since the invitation to Monash increased the probability of his going there but did not increase the probability of his going to Oxford. Similarly, the favoring criterion rightly excludes the explanation by appeal to Lewis's desire to go to a place where he has good friends, since while that raises the probability of his going to Monash, it also raises the probability of his going to Oxford.

An account of contrastive explanation in terms of shifting probabilities appears naturally to capture a notion of a cause that marks a difference. Moreover, as Christopher Hitchcock (1999) observes, it also provides an account that appears to capture the independently plausible link I flagged above between contrast and presupposition, namely that to explain P rather than Q is to explain why P, given P or Q. Nevertheless, the favoring account seems too permissive. Suppose we ask not why Jones rather than Smith contracted paresis, but why Jones contracted paresis rather than remaining relatively healthy. I take it that the fact that Jones had syphilis does not explain this contrast concerning Jones alone, since he might well have had syphilis even if he had been relatively healthy, given that so few with syphilis go on to contract paresis. Yet Jones's syphilis meets the favoring condition in this case, since it raises the probability of Jones with paresis without raising the probability of healthy Jones. (Lewis's counterfactual account gives the same wrong answer here, since the syphilis would not have been a cause of Jones remaining healthy.) Moreover, a favoring account faces the same difficulty that Lewis has already faced: it wrongly rules in the use of the invitation to Monash to explain why Lewis went to Monash rather than to Oxford, even when Oxford also invites. For even in such a case, the invitation to Monash raises the probability of going to Monash without raising the probability of going to Oxford. (Eric Barnes has made this objection to Hitchcock's version of a favoring account in correspondence; for Hitchcock's reply see his 1999: 605–6.) What seems to be wrong with requiring only that the cause of P would not have been a cause of Q, or with requiring only that the cause of P not raise the probability of Q is that neither of these accounts captures the need in contrastive explanation not just for the presence, for example, of an invitation to Monash, but also the need for the absence of an invitation to Oxford. To explain why P rather than Q, we seem to need not just a cause of P, but also the *absence* of a corresponding event.

Causal triangulation

In an attempt to improve on the counterfactual and favoring approaches to contrastive explanation, consider John Stuart Mill's Method of Difference, his version of the controlled experiment, which we discussed in chapter 1 (Mill 1904: III.VIII.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 (Garfinkel 1981: 40). Accordingly, I propose that, for the causal explanations of events, explanatory contrasts select causes by means of 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 case 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, or a single cause with differential probabilistic effect, as a favoring account suggests, contrastive questions select as explanatory an actual causal difference between P and not-Q, consisting of both a presence and an absence. 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 led 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 is 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. As most of these examples illustrate, the required absence in the case of not-Q is typically an absence from the causal history of not-Q, but this is not always the case. Where both Jones and Doe have syphilis, but only Jones

also has paresis, Jones's syphilis clearly does not explain why he rather than Doe has paresis; nevertheless, Doe's syphilis is not a cause of Doe not having paresis. The Difference Condition must thus be read as requiring not just the absence of the corresponding event from the causal history of not-Q, but its absence, *tout court*. (The need to rule out cases where the corresponding event is present but not part of the causal history was made clear to me by Michael Gaylard and Tom Grimes.)

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 applies to incompatible contrasts too. 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. Similarly, the invitation from Monash explains why Lewis went there rather than to Oxford and satisfies the Difference Condition, so long as Oxford did not invite. The 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 case of not-Q where, roughly speaking, a corresponding event is something that would bear the same relation to Q as the cause of P bears to P.

The application of the Difference Condition is perhaps most difficult to see in cases where the contrastive question does not supply two distinct instances, like Jones and Smith. For example, we may explain why a particle was deflected upward rather than moving in a straight line by observing that the particle passed through a particular field: this field is a causal presence that explains the contrast, but it is not clear in such a case what the corresponding absence might be. (I owe this point and example to Jonathan Vogel.) This sort of case is not unusual, since we often ask why a change occurred rather than the status quo. Similarly, instead of asking why Jones rather than Smith contracted paresis, a case to which the Difference Condition is easy to apply, we might ask a similar question about Jones alone, that is why he contracted paresis rather than staying healthy. As in the particle case, we seem here to have only one instance, and that seems at first to block the application of the notion of a corresponding event.

In fact I think the Difference Condition does apply in single instance cases, including the case of the wayward particle, though this is perhaps particularly difficult to see at first because what the Condition requires in that particular case is the absence of an absence. To see how the Condition works in such cases, it is helpful to work up to it in stages (which will also give us another good illustration of how explanation is sensitive to contrast).

Suppose first that there are two particles, one with a field and deflection upward and the other with no field and no deflection. This is an easy case for the Difference Condition: the presence of the field in the one case and its absence in the other explains why one particle rather than the other was deflected, because the field in the one case is a cause of the deflection, and the corresponding event in the other case would be a similar field there, which is duly absent. And even in a single particle case, there need be no difficulty. Thus, if the question is why the particle deflected upward rather than downward, we can explain this in terms of the presence of a field with a particular orientation, since the corresponding event would be a field oriented in the opposite direction, again duly absent. Now we can return to the original example, where we ask why the particle was deflected upward rather than moving in a straight line. The Difference Condition again applies after all, since we have the presence of the field, where in this case the corresponding event would be the absence of any field, and this (absence) is absent, there being a field present. Although the application of the Difference Condition is easiest to see in cases like those where the Method of Difference applies, with two quite distinct instances in one of which the effect occurs and in the other of which it does not, the Condition applies to single instance contrasts as well.

   I hope this helps to show how to apply the notion of a corresponding event in cases where there is only one instance. There is, however, another related challenge to the notion of a corresponding event: not that it is sometimes inapplicable, but that it is vague. I have no full response to this difficulty, but will worry it a bit here. As a first approximation, I suggested above that a corresponding event is something that would bear the same relation to Q as the cause of P bears to P. But what does this mean? Clearly it cannot mean simply anything that would have caused Q (Achinstein 1992: 353). If Lewis was invited to both places, the invitation to Monash does not explain why he went there rather than to Oxford, even though of course he was not abducted to Oxford. We would do better to think of the Difference Condition as requiring the presence of one token of a type and the absence of another token of the same type. But not any type will do. Thus, if both Monash and Oxford invited Lewis, the invitation to Monash will not explain the contrast, even if that invitation falls under the type 'invitation printed on pink paper' and the invitation to Oxford is not of that type. Perhaps this difficulty can be met by requiring that the type be pitched at a level of causal efficacy. Thus the type 'invitation' is appropriate because it is in virtue of being an invitation that the token caused Lewis to go to Monash. So it is only the absence of an invitation to Oxford, not merely the absence of an invitation on pink paper, that would allow the invitation to Monash to explain the contrast. (On the other hand, if Lewis had been the sort of person more likely to be swayed by invitations on pink paper, then that type would be explanatorily relevant.)

The idea of the presence of one token and the absence of another is, however, still only an approximation to the requirements of the Difference Condition. For one thing, the absence needs to be tied to the foil. (The absence of an invitation to Cambridge will obviously not help us to explain why Lewis went to Monash rather than to Oxford.) For another, although many explanatory presence–absence pairs are naturally seen as two tokens of the same type, not all are. Thus if you ask why the mercury in the thermometer rose rather than fell, one token is of increasing temperature, while the other is of decreasing temperature. I have attempted to capture both of these points in the gloss I gave above on the corresponding event for the P/Q contrast – 'something that would bear the same relation to Q as the cause of P bears to P' – but the account does remain somewhat vague, and it is not clear how to make it more precise. It is for example now tempting to return to the condition that the corresponding event must be something that would have been a cause of Q, if Q had occurred, only now as an additional requirement alongside something like the token-type condition. But while satisfying this further condition may yield particularly satisfying contrastive explanations, it is not necessary. For we can explain why Kate rather than Frank won the prize by pointing out that she wrote the better essay, even though had Frank's essay been better than Kate's, that still would not have assured him of the prize, since a third party might have written something better still.

I have been unable to give a precise account of the notion of a corresponding event. By way of mitigation, I would say that any unclarity in the notion of a corresponding event is one that we clearly negotiate successfully in practice, since it is a notion we deploy frequently and uncontroversially in inferential contexts when we apply Mill's Method of Difference to infer causes from effects and their absence. As Mill puts it:

> If an instance in which the phenomenon under investigation occurs, and an instance in which it does not occur, have every circumstance in common save one, that one occurring only in the former; the circumstance in which alone the two instances differ is the effect, or the cause, or an indispensable part of the cause, of the phenomenon. (1904: III.VIII.2)

Here 'the circumstance in which ... the two instances differ' is tantamount to the presence–absence pair in contrastive explanation, and faces the same vagueness problems that I have been worrying in the notion of corresponding event. How do we tell whether a circumstance attaches to an instance? And do two instances differ if they both for example contain invitations but only one is issued on pink paper? A fuller analysis of the notion of differing circumstances would be welcome, but the application and development of the Method of Difference does not wait on this, and I suggest that the

situation is similar with respect to the notion of corresponding event and the Difference Condition on contrastive explanation.

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 specialties 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 chapter 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. The latter question is not part of providing a model of explanation, as that task has been traditionally construed. It is no criticism of the deductive-nomological model that it does not tell us which phenomena we care to explain, so long as it tells us what counts as an adequate explanation of the phenomena we select; similarly, it is no criticism of my account of contrastive explanation that it does not tell us why we are interested in explaining some contrasts rather than others. Still, an account of the considerations that govern our choice of why-questions ought to form a part of a full model of our explanatory practices, and it is to the credit of the contrastive analysis that it lends itself to this. As we will see in later chapters, our choice of why-questions is often governed by our *inferential* interests, so that we choose contrasts that help us to determine which of competing explanatory hypotheses is correct. For now, however, we may just note that 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 (cf. Barnes 1994). When the histories are too 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.

   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 more detailed comparison of my analysis with the deductive-nomological model. When we ask contrastive why-questions, we choose our foils to point towards the sorts 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. This focuses our inquiry on causes that will illuminate the reason our expectation went wrong. 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 behavior, 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. More than one of these may be explanatory: my account 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, and one that I have already mentioned, is that when they are available we usually prefer explanations where the foil would have occurred if the corresponding event 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 specialties do not fit the department's needs. Suppose also that, had Baker's specialties been appropriate, he would have gotten the job, even without a letter from Quine. In this case, the difference in specialties is a better explanation than the difference in letters. As we have seen, however, 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. To give another example, 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 (but cf. Carroll 1999). 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' over-determination. 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. Again, we sometimes don't 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 tend to prefer them.

Another reason satisfying the Difference Condition is not always sufficient for a good contrastive explanation is that a difference favoring the fact may be balanced against another favoring the foil. If I tell you that Lewis went to Monash rather than Oxford because only Monash invited him, you might reply, 'Yes, but Oxford has much better book shops, and Lewis loves book shops.' In such a case, I will have to supplement my original explanation by showing, or at least claiming, that the actual cause of the fact trumped or outweighed the potential cause of the foil. Thus I might claim that his preference for places that invite him was stronger than his preference for places with outstanding book shops. Of course this might not be true: the difference I originally cite may not by itself be stronger than the countervailing force you mention. In this case, I must find other or additional differences that are. Here the hardware of a probabilistic favoring approach has a natural application. There are doubtless other principles that also play a role in determining which 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 seem to 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.

How does the Difference Condition behave for these 'global' contrasts? I once thought the answer was 'pathologically', since the Condition would impossibly require that we find a cause of P that is at once present and absent; but I now see (thanks to Naomi Rosenberg) that things are not so bad. When the question is, 'Why P rather than not-P?', what the Difference Condition requires is the absence of something that bears the same relation to *not*-P that the cited cause bears to P. If C is the cause of P, then what would bear the same relation to not-P is presumably not C itself, but something else. But what would it be? The difficulty in answering the question arises because 'not-P' is not a limited negation, but encompasses all the different specific ways P might not have been the case. The 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'. 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 we do not, however, 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. I nevertheless remain unsure whether every apparently non-contrastive question should be analyzed in contrastive form, so I am an agnostic about the existence of non-contrastive why-questions.

Finally, let us compare my analysis of contrastive explanation to the deductive-nomological model. First, as we have already seen, a causal view of explanation has the merit of avoiding all the counterexamples 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 necessary 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, at 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, among other things, a symptom of 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 view of explanation over the deductive-nomological model. One odd feature of the 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 (assuming they are not themselves laws). 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 contrastive model avoids this problem. Another related and unhappy feature of the deductive-nomological model is that, as we have seen, 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 view 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 might 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 scientific explanations, unlike ordinary explanations, do not exhibit the interest relativity of foil variation that a contrastive analysis exploits, so a contrastive analysis does not apply to scientific explanation (Worrall 1984: 76–7). 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 shows, even a full deductive-nomological explanation may exhibit interest relativity: it 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 orbit, 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 of why members of a species possess a certain trait may explain why they have that trait rather than various maladaptive traits, but it may not explain why they have 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 an appeal to the evolutionary history of the species, insofar 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 view does not seem to account for this. For even if the fact to be explained carries no restricting contrast, the contrastive view, if it is extended to this case by analyzing 'Why P?' as 'Why P rather than not-P', requires at most that we cite a condition that is causally sufficient for the fact, not that we actually give any laws. In reply, one might mention first that laws may nevertheless be part of a correct analysis of the causal relation itself, and that knowledge of laws is sometimes essential evidence for causal claims. Moreover, the contrastive view can help to account for the explicit 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 (Friedman 1974: 5; though explaining particular events is also important when, for example, scientists test their theories, since observations are of particular events). Something like 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 here we need some general cause (cf. Lewis 1986: 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 view, 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 view 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 between the scientific emphasis on theory and the contrastive view of explanation 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 will provide many explanations. The contrastive view 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 view 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 behavior under diverse conditions. So I conclude that the contrastive view of explanation does not fly in the face of scientific practice.

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, among other things, not all explanations are causal and since the choice of foil is not the only factor that affects the appropriate choice of cause. The model 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 model. 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 cause 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 even close to being sufficient for P, so long as it is part of a causal difference between P and not-Q. Causal triangulation also elucidates 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, the interest relativity of explanations reduces to the important but unsurprising point that different people are interested in explaining different phenomena. Moreover, the Difference Condition shows that different interests do not require incompatible explanations to satisfy them, only different but compatible causes. The mechanism of causal triangulation also helps to account for the failure of various attempts to reduce contrastive questions to some 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. Lastly, we will find that the structure of contrastive explanations helps us with the problem of describing our inferential practices, a problem whose difficulties we met in chapter 1, when it is wed to Inference to the Best Explanation, an account to which we now finally turn.