

Knowledge and Explanation

1 Introduction

Craig (1990) casts doubt upon the project of trying to give the traditional sort of necessary and sufficient conditions for *A knows that p*. He interprets the inadequacy of existing analyses of knowledge as evidence that our concept of knowledge is complex and diffuse, and concludes that we should aim to understand it by thinking about the rôle the concept plays in our lives, rather than by trying to find necessary and sufficient conditions for the truth of knowledge ascriptions.

There is surely something right about Craig's view: we are unlikely to succeed in any attempt to analyse away the intricacies in our concept of knowledge. We cannot realistically hope to uncover a set of necessary and sufficient conditions for *A knows that p* which are in all cases either clearly satisfied or clearly not satisfied. Nor, I suspect, is it possible to offer necessary and sufficient conditions for knowledge which are widely accepted as being more securely understood than knowledge itself.

But to conclude that we should therefore stop trying to find *any* necessary and sufficient conditions for *A knows that p* may be going a step too far. Perhaps we can further our understanding of what knowledge is by uncovering necessary and sufficient conditions which *share* the complexities of the target proposition. We can aim, not to "clean up" these problematic features, but to capture them in

ways which may prove illuminating, even if they do not amount to reductive analyses.

In this paper I attempt a project of this kind. I propose a necessary and sufficient condition for *A knows that p* which is, although recognizably similar to the traditional sets of conditions, arguably immune to the kind of counterexample which tends to deter philosophers from thinking that any illuminating conditions can be found.ⁱ I present this condition, however, not as an analysis of knowledge, but rather as a way of getting a handle on the concept and furthering the effort to understand what its rôle in our lives might be. Taken in this spirit, the current proposal is not at odds with the principles that motivate Craig's view.

In denying my proposal the status of a reductive analysis, I am mindful of the fact that it will tell us little more than that knowledge is 'non-accidental true belief'. What it offers is a (hopefully fruitful) way of spelling out what is meant by 'non-accidental' in this context. In what follows, I shall write 'KAp' for 'A knows that p' and 'BAp' for 'A believes that p.' I shall propose that KAp just in case BAp and it can be said (under specific circumstances, to be described shortly) that A believes p *because* p is true. But this is not a causal account of knowledge. The 'because' signals not causation, but explanation.

Explanationist accounts of knowledge somewhat similar to mine have been proposed by Alan Goldmanⁱⁱ (1988 and elsewhere) and Steven Rieber (1998). However, neither of these philosophers exploits what I take to be the most important strengths of an account of this type. Consequently Goldman's and Rieber's accounts face serious objections, which I think may be avoided by a

sufficiently sophisticated explanationism. These objections will be discussed in §2 below.

Neta (2002) has offered an account of knowledge which he claims is a modification of Rieber's. In fact, however, it is significantly different: Neta proposes that KAp iff p is 'the reason for' BAp (668). This sounds like an explanationist account, until we discover that 'reasons' are to be contrasted with 'non-rational causes', revealing that this is essentially a justified true belief account in the traditional style. (The proposal is that KAp iff A reasonably believes that p on the strength of A 's conclusive evidence that p (668). While explanationist accounts of the kind I am interested in do in some sense focus on 'reasons' for belief, it is not required on these accounts that they be epistemic reasons, or even reasons for the subject at all.)

My own proposal is that KAp just in case the following condition is met:

K: p is a good explanation of BAp for someone not acquainted with the particular details of A 's situation (an 'outsider').ⁱⁱⁱ

I shall say a bit more about outsiders in a moment. First note that, with **K** in place, we do not need the usual clauses stating that p and BAp are true, since only true propositions make good explanations, and only true propositions can be explained. Note also that **K** does not require that p be *the* good explanation of BAp (or even the good explanation for an outsider of BAp). We respect the fact that explanations may be more or less good depending on the interests and

intentions of their recipients, which will vary from context to context. All **K** requires is that there be *some* context in which an outsider might ask ‘Why BAp?’ and be well-answered by ‘Because p’.

An outsider, in the sense envisaged here, is a person O who meets three conditions:

- (a) O is rational, and can understand the content of A’s belief that p (i.e. is capable of entertaining the proposition p).
- (b) O is aware of *commonplace* facts about people and their mental lives, i.e. facts about what it is like, in general terms, to be a rational thinking person.
- (c) O is not aware of any *special* facts about A or A’s situation. O is aware that A is a person and that A believes that p, but that is all.

In the remainder of this paper, I shall motivate my condition **K** by showing how it avoids Gettier counterexamples (§§3-4) and how it accommodates the motivations for various other theories of, and intuitions about, knowledge (§5). I shall then respond to an objection in §6. First, however, I must show how my account differs from Goldman’s and Rieber’s, and why some improvement on their accounts seems necessary. This will occupy §2.

2 Other Explanationist Accounts

Goldman (1988, 22) proposes that KAp iff p ‘enters prominently into the best explanation of BAp’, and Rieber (1998, 194) suggests that KAp iff p ‘explains’ BAp.

Let us consider Goldman's account first. I do not object to this account on the grounds that it appeals to 'the best explanation' of BAp, rather than (as in my condition **K**) 'a good explanation to an outsider'. In fact this is something of a red herring, as it is later undercut by disclaimers. For instance, at 25-6 we are told that 'the explanation for a belief relevant to its evaluation as a claim to knowledge may not be the explanation uniquely best for all explanatory purposes'.

Goldman's response to a certain kind of objection might be criticized, but this problem is not devastating either. The objection in question states that facts about the future cannot enter into the best explanations of current beliefs, so that on Goldman's account we cannot know such facts. In fact, I see no reason to doubt that facts about the future explain certain facts about the present, including facts about present beliefs, so the objection does not seem worrying to me. Goldman, however, feels that some more substantial response is needed (perhaps because he has a chance-raising analysis of explanation, and tacitly envisages chance-raising as causal or quasi-causal). He therefore weakens his account so as to allow cases where some common fact enters into the explanations of both p and BAp to count as cases of knowledge. He allows, for instance, that KAp whenever the existence of some evidence e explains both BAp and p.

This, of course, is isomorphic to the attempt made by Alvin Goldman (1967, 364ff) to rescue his causal theory of knowledge by allowing that there is knowledge wherever a single event causes both p and BAp. And the strategy is problematic in both cases for the same reason: the resulting accounts seem to be

too weak. Any pair of propositions p and $B\Delta p$ have *some* part of their explanatory (and causal) histories in common, if only the big bang. So in particular, any pair of propositions p and $B\Delta p$ will always have some part of their explanatory (causal) histories in common. So any belief counts as knowledge.

Can this sort of worry be finessed by demanding that the common element be something ‘prominent’ in the ‘best’ explanations of both p and $B\Delta p$? This is not clear; for one thing, the resulting account is still too weak by Goldman’s own lights. For if we accepted it, we ought also to accept that, on buying a lottery ticket, I know it will not win. After all, the existence of a great many other tickets enters prominently into the best explanations both of the fact that my ticket will not win and of the fact that I believe it will not win. But Goldman thinks that I do not know my ticket will not win (52).

This is not a conclusive objection, however; some may be willing to accept that I know my ticket will not win. And anyway, as I have said, the objection which prompts Goldman to modify his account in this way is not really worrying. So even if he has failed to address it properly, this does not show there is anything fundamentally wrong with his account.

What *does* seem to be wrong with the account is that, for reasons independent of the amendment to allow for future knowledge, it is too weak as it stands. For p might ‘enter prominently’ into the ‘best’ explanation of $B\Delta p$ in all sorts of cases where we would not want to ascribe knowledge. For instance, a bang on the head might, simply through having a certain kind of physical impact on my brain, and not via any rational process, cause me to believe I’ve just

received a bang on the head. In such a case, Goldman's condition is surely met: a prominent part of the best explanation of my belief is the fact believed. Yet I do not *know* that I have just received a bang on the head.

To deal with such problems, Goldman complicates his account by imposing an additional constraint, which he calls 'Condition C'. In order to understand it we must note that, for Goldman, to explain something is to raise its probability. Condition C reads as follows:

The explanans for a belief, which explanans explains (type 1) or is explained by (type 2) the fact to which the belief refers, must be such that: if it is a type 1 explanans, its raising the probability of the belief is itself made more probable by its raising the probability of the fact to which the belief refers; if it is a type 2 explanans, its raising the probability of the belief is made more probable by its being made more probable by the fact to which the belief refers. (36)

(Note that the 'type 2' explanation is the normal case, while the 'type 1' explanation is the kind allowed by Goldman's amendment to deal with knowledge of future facts.)

Unfortunately, even when condition C is incorporated, Goldman's account is too weak. It is susceptible to counterexamples of the sort described by Gettier (1963). Suppose I see Professor Craig driving a BMW, and conclude that one of the Cambridge Philosophy professors owns a BMW. Unbeknownst to me, the BMW Professor Craig is driving is owned by Professor Heal, who has lent him the car for the day. In this situation, we do not want to say that I *know* one of the Cambridge Philosophy professors owns a BMW.

Let A be me, let $p = \textit{one of the Cambridge philosophy professors owns a BMW}$, and let $q = \textit{Professor Craig is driving a BMW}$. Now q surely features prominently in ‘the best’ explanation of BAp, and p surely features prominently in ‘the best’ explanation of q . Thus (presumably) p features prominently in ‘the best’ ultimate explanation of BAp, so that Goldman’s original condition on knowledge is met. What’s more, condition C is also met: q is a type 2 explanans of BAp, and [the fact that p makes q more probable] makes it more probable that [q makes BAp more probable]. This is because I am a rational person, and so the weight I attach to evidence like q in deciding whether to believe p is sensitive to the probabilistic relations between q and p in the way C demands. Indeed, Condition C is best thought of as a kind of rationality constraint, enforcing this kind of sensitivity. Viewing it as such makes plain why it is susceptible to Gettierization, since what Gettier cases show is that beliefs can be perfectly true, justified and rational, and yet still fail to amount to knowledge.

Goldman’s questionable analysis of explanation is another weak point in his account. As I mentioned, it is a chance-raising analysis. More precisely, Goldman demands that an explanans must raise the chance of the explanandum, without there being some other background factor which wholly accounts for this difference (23). The trouble is that sometimes one proposition can explain another without raising its probability. Consider, for instance, ‘failsafe’ cases such as the following. There is a 100% chance that my alarm clock will go off at 7.30am and wake me up. But, independently of anything to do with my alarm clock, there is also a 100% chance that my cat will jump onto the bed at 7.31am

in a way that would be sufficient to wake me up even if the alarm clock hadn't already done so. So whether or not the alarm goes off, there is a 100% chance that I will be awake by 7.32am. Suppose the alarm does in fact go off. Then, the fact that the alarm goes off supplies a perfectly good explanation of why I am awake at 7.32, even though the alarm clock's going off did not raise the chance of my being awake at 7.32, which was already at 100%..

There are ways of responding to this point by appealing, not to the chance of the explanandum's being true if its explanans is not true, but to the chances of its being brought about *in the same way* if the explanans is not true (a manoeuvre that parallels the familiar Nozickian appeal to the 'methods' by which a belief is formed.) Whether or not any mileage in this line of response depends on whether there is a stable notion of *being brought about in the same way* which is fine-grained enough to enable us to answer the objection without being so fine-grained that anything involved in the causal history of p will trivially count as explanatory of p (because if it had not happened p could not have happened in *exactly* the same way).

Unlike Goldman, I do not depend on any particular account of explanation in this paper. I shall rely on intuitions about explanation, such as the view that an explanation is an answer to a why-question, and the view that why-questions are implicitly contrastive (these claims are defended by e.g. van Fraassen 1980 and Garfinkel 1981). In defence of this strategy, I would say that these intuitions are so fundamental they ought to be respected by any account of explanation.

Rieber (1998) proposes an account of knowledge belonging to the same research programme as Goldman's and mine. He argues that KAp iff p 'explains' BAp . The main problem with this is that it is vague as it stands; one could interpret it as identical to the condition \mathbf{K} which I propose, or as identical to the condition proposed by Goldman, or in various other ways. But in his note 7, Rieber states that Goldman's condition is sufficient for knowledge. This shows that, on his intended reading, p 'explains' BAp whenever p enters prominently into the best explanation of BAp . For this reason Rieber's account, like Goldman's, is too weak.

We can see this by considering one of Rieber's own examples (201-2):

An apple falls on Newton's head. By an amazing coincidence, Newton has a brain lesion such that, whatever rate the apple falls, the impact of the apple will (together with the lesion) cause Newton to believe that objects fall at precisely this rate. As it happens, the apple falls at 32ft./sec.^2 , and Newton walks away with the belief that objects fall at this rate.

Rieber rightly says that this is not a case of knowledge, and he also claims that his condition on knowledge is not met in this case, since explaining Newton's belief by citing the fact believed "leaves out the most salient and important part of the explanation, namely that Newton has a very peculiar brain lesion." But is he entitled to say that this means his condition is not met? It seems not. His condition does not explicitly demand that p be an explanation of BAp which cites the 'most salient and important facts'. Although doing so required of a *good* explanation, and we might be moved charitably to interpret Rieber's condition as

demanding that *p* be a *good* explanation of BAp, such an interpretation would be mistaken. Rieber explicitly does *not* appeal to explanatory goodness in his conditions for knowledge, because he thinks that goodness comes by degrees and knowledge does not (195).^{iv}

Perhaps, however, he could say that, because it leaves out salient facts, the claim that objects fall at 32ft./sec.² simply does not ‘explain’ Newton’s belief. But then Rieber owes us some account of when enough ‘salient’ facts have been included for one proposition to count as ‘explaining’ another. And besides, as we saw, his endorsement of Goldman’s condition as sufficient reveals that he thinks it is enough for *p* to ‘explain’ BAp if *p* features prominently in the best explanation of BAp. It can do this without by itself supplying all the salient facts.

Neta (2002, 667-8) has furthermore argued that the sort of response Rieber gives to the Newton problem does not resolve the underlying difficulties. I shall postpone discussion of Neta’s argument, because it can also be construed as an objection to my own proposal, and will be considered as such below. But as we shall see, Neta’s argument is spurious. What Rieber says about the Newton case is exactly the right thing to say (though it is possible to be more explicit as to why this is so). It’s just that Rieber does not really seem to be entitled to say it.

3 Gettier-Immunity

I now return to my own proposal, whereby a belief that *p* counts as knowledge iff it can be well explained to an outsider just by citing *p*.

An important feature of this proposal is its ability to deal with Gettier-type cases: cases where a subject has a true belief which meets various supposed sufficient conditions for knowledge and yet does not amount to knowledge. It will be significant if a condition on knowledge can be shown to be immune to counter-examples of this kind, which have floored so many other proposals.

Whenever we are given a very simple explanation of something, we assume this is because all other relevant facts are as we would expect them to be. Suppose I explain my going out by telling you I want some chips. This is an adequate explanation if all other factors are normal (e.g. I believe the chip shop is open and I can get some chips there). But suppose I believe that all the shops are closed, and that going for a brisk walk will take my mind off my chip craving. Then explaining my going out just by saying I want some chips seems inadequate and misleading. Although my desire for chips will still form *part* of an explanation of my behaviour, it is no longer a good explanation *by itself*. That is, not unless you are already closely acquainted with my customary methods of dealing with unwanted cravings.

It will be helpful in this context to talk of *collapsible* and *non-collapsible* explanations. Suppose I explain my going out to an outsider (someone not acquainted with any specific details about me) by saying I want some chips, I believe that the chip shop is open and I believe that I can buy chips there. This explanation is *collapsible*: it would not mislead my audience if I left out the fact that I believe the chip shop is open and that I can buy chips there, and collapsed the explanation into a much simpler one, citing only my desire for

chips. By contrast, suppose I explain my going out to an outsider by saying I want chips, but I don't want to want chips, and believe that by going for a walk I can take my mind off my chip craving. This explanation is *not collapsible*: if I collapsed it into a simple explanation citing only my desire for chips, this would be misleading. The outsider would assume I was going out in order to *get* some chips. By contrast, if I were talking to someone I knew very well, the latter explanation *might* be collapsible: such a person would be likely to assume I was going out to try and forget about the desire for chips, knowing that this is what I usually do when I have that sort of desire.

The rôle of the outsider in my account of knowledge is to set a high standard for collapsibility. The less my audience knows about me and my circumstances, the less likely it is that explanations of my behaviour can be collapsed without being misleading. The outsider, who knows nothing particular about me, demands an uncollapsed explanation if anyone does. Hence it is *difficult* for a belief to count as knowledge by the lights of condition **K**. This is what makes it preferable to Goldman's account, which (as we saw above) sets rather looser standards on knowledge, and is thus too weak.

My strategy in this section will be to argue that Gettier cases are cases where there is *some* explanatory connection between *p* and BAp, but where that connection is unusual in a way that must be mentioned when giving an explanation of BAp to an outsider. This means that to explain BAp to an outsider by citing *p* alone will be inadequate in such cases, because it will be misleading. Thus condition **K** is not met in Gettier cases.

To see how this is supposed to work, consider the following example. Suppose, as before, that I see Professor Craig driving a BMW, and as a result believe that one of the Cambridge philosophy professors has bought a BMW. As before, the BMW Professor Craig is driving has been lent to him by Professor Heal. Now, I believe that one of the Cambridge philosophy professors has bought a BMW, and what's more the fact that one of them *has* bought a BMW can figure in a good explanation of that belief to an outsider. Yet, we feel, I do not *know* that one of the Cambridge philosophy professors has bought a BMW.

This is because condition **K** is not met. In this unusual kind of case, we would need to say *more* to give a good explanation of my belief to an outsider than that one of the Cambridge philosophy professors has bought a BMW. If the outsider were given this simple explanation, he would assume that the professor who bought the BMW was the same as the one I think bought it (at least, if I have any belief as to which of them that is). But this is not the case; I think Craig bought the BMW, though in fact it was Heal. So the simple explanation is inadequate because it is misleading on this point. There is an explanatory connection between p and BAp here, but not one which is collapsible when talking to an outsider.

It should be clear that this strategy generalizes, allowing us to deal with any other Gettier-type case we care to contrive. The 'deviancy' of the link between p and BAp, which makes Gettier cases so problematic for other accounts of knowledge, is just what supplies us with a reason to say condition **K** is not satisfied in these cases.

Gettier cases are cases of justified but accidental true belief. We should note that the account I am defending also deals with cases of *un*justified accidental true belief. Suppose I am hit on the head with a hammer, and this causes a brain injury, one effect of which is that I lose my memory completely but irrationally start believing I have just been hit on the head with a hammer. In this situation one cannot explain my belief to an outsider just by citing the fact that I have been hit on the head with a hammer, since this would lead the outsider to think that my belief was rational (e.g. that I remembered being hit). So **K** is not met in this sort of situation either, and the problem faced by Goldman (see 5-6 above) is avoided by my account.

In general, the account makes sense of the fact that we intuitively take justification to be necessary for knowledge. If A's belief that *p* is formed in an irrational way (without justification), then a simple explanation of BAp which cites only *p* will always be misleading to an outsider. Irrational methods of belief-formation are unusual, so the outsider will assume that nothing of that kind is going on unless she is told otherwise. Hence in cases of irrational or unjustified belief, condition **K** is not met, and we do not ascribe knowledge.

It may be objected that I seem to be suggesting nothing more than that to know that *p* is to have a belief that *p* which is acquired because of *p* and 'in the usual kind of way'. For I seem to be claiming that *p* is an adequate explanation of BAp just in case there is an explanatory link between *p* and BAp which involves nothing "unusual" (i.e. nothing which would need to be mentioned to someone who expects things to be as they usually are).

In reply, I should repeat that, essentially, I *agree* that what I am proposing is merely a development of the thought that knowledge is true belief acquired in the ‘normal’ sort of way; I do not presume to have analysed away the notion of normality which informs our concept of knowledge. What I do claim for my proposal is that it offers us a useful way of *dramatizing* this notion, and thus helps us to understand it a little better. I am trying to reduce the problem to a familiar one, namely that of saying when an explanation is adequate in a particular context. (Without offering a full account of explanation, I cannot say much more about the intuition that, in Gettier cases, *p* is *not* an adequate explanation of BAp to an outsider. But I trust that intuition is strong enough not to need support from any particular theory of explanation.)

The proposal certainly doesn’t seem to be rendered inert by dint of being a development of the claim that knowledge is normal true belief. It can, for instance, help us understand why not *all* kinds of explanatory abnormality are relevant to knowledge ascriptions, which mere handwaving at ‘normal true belief’ cannot do. Suppose a wartime diary is preserved despite the extreme unlikelihood of this happening. We can colour the story with tales of burning libraries and courageous escapees bearing concealed manuscripts. The diary contains a genuine account of its author’s attempt to overthrow the current military regime, which was unsuccessful, and was so well concealed that this diary is the only remaining evidence for it. The diary is then published and I read it. Do I know that the author tried to overthrow the regime (= *p*)? We surely want to say I do. But don’t we have to mention the unlikely escape and the

burning libraries in explaining my belief to an outsider? They are hardly normal circumstances. So, surely, explaining my belief that p to the outsider just by citing p is inadequate, implying as it does that all factors not mentioned in the explanation are as the outsider would expect.

In fact, however, we do not need to cite these unusual factors to explain my belief that p. The reason is that, notwithstanding the unusual *features* of the explanatory connection between the truth of p and my belief that p in this case, the connection *itself* is of an utterly mundane kind: I come to believe p through reading the honest testimony of someone with first-hand experience of the relevant facts. By contrast, in Gettier cases, *the very connection* between p and BAp is of an unusual kind. Seeing Professor Craig driving a BMW is not a normal connection to the fact that one of the Cambridge philosophy professors has bought a BMW, when the person who has done so is Professor Heal.

It is a general fact about explanations that unexpected *features* of explanatory links do not need to be mentioned to avoid misleading one's audience, although unexpected *links* do. Suppose I explain why the light came on (p) by saying that I flipped the switch (q). This will be inadequate, if what actually happened was that Anna saw me flipping a switch which she knew did not work and therefore kindly decided to turn the light on herself using another (functional) switch. The explanatory link between p and q is not a normal one, so we must say more than p to explain q (unless our audience already knows a lot about the situation under consideration).

By comparison, this explanation (q) of this fact (p) *would* be adequate if all that was odd about the case was that the switch was connected to the light insecurely, by forty-year-old wiring. For in this case, although the explanatory connection between p and q would have unusual *features* (the insecurity of the connection and the age of the wiring), it would be typical of the *kind* of connection you normally get between switches being flicked and lights coming on.

When knowledge is understood as I propose, this general fact about explanations enables us to understand the impulse to say that in the case of the wartime diary I have knowledge, whereas in the case of Professor Heal's BMW I do not. We understand which sorts of abnormality are relevant to knowledge ascriptions by understanding which sorts of abnormality are relevant to explanation.

My definition of an outsider (see 3 above) has been doing crucial work in this section. Before we move on, therefore, we might take a moment to consider whether that definition is really correct. What kinds of 'commonplace facts' is the outsider aware of? If we give her too much, so as to include (e.g.) familiarity with typical human epistemic failings, we may end up allowing something to count as a case of knowledge where in fact it is only the output of a such an epistemic failing. If the outsider is aware that we humans typically conclude that p in circumstances C whether or not it is justified, and that whenever p is true circumstances C obtain, then explaining BAp by p may be adequate for that

outsider, even when it is only by lucky chance that A's belief is in fact true in this case. For citing p will be adequate to explain to her why C obtains, and this will explain to her why BAp . This is so even if (say) only five percent of cases where C obtains are cases where p , and A indiscriminately believes p whenever C obtains.

To avoid this sort of problem, the outsider must know nothing except the bare bones of what it is to be a thinking person. But there seems to be a risk that if she knows so little, condition **K** will become too strong. Suppose aliens who do not share our sensory modalities can count as outsiders by our definition.^y Then, it seems, we may end up ruling out many perfectly ordinary explanations of beliefs of the kind we want to include, simply because they wouldn't be good explanations for these aliens. For instance, an alien from a world whose inhabitants lacked anything resembling a sense of sight (or other means of responding to the frequency of light waves) presumably would not find 'Snow is white' a good explanation of my belief that snow is white. But we don't want to say for this reason that I do not know that snow is white. Do we need, then, to specify that aliens of this sort cannot count as outsiders? If so, how many other kinds of rational being should be excluded? Could someone from a very different terrestrial cultural background fail to be an outsider?

Actually, I don't think we need be worried by this line of thought. It may be true that an alien who had no familiarity with visual perception would not find 'Snow is white' a good explanation of my belief that snow is white, but we need to consider *why* this is so. If the problem is simply that they would not

understand the explanation, these aliens are already ruled out as potential outsiders by dint of clause (a) on p. 3 above. On the other hand, if they *can* understand ‘Snow is white’, why should we doubt that they will find it a good explanation of my belief that snow is white? Even if they have no idea how exactly I managed to come by that belief, they will not be misled by the explanation: presented with it, they will assume, correctly, that I used some sensible method of investigating the world around me and responded appropriately to the deliverances of that method. The definition of an outsider offered in §1 above is adequate for our purposes.

4 Neta’s Objection

I shall now take up the postponed discussion of Neta’s objection, which is directed against Rieber’s attempt to give a certain response to the problem case of Newton’s apple. What Rieber says about this case is similar to what I have said about cases of irrational true belief where there is some explanatory link between the belief and the fact believed. Recall that Rieber claims that in the Newton’s apple case the explanation of BAp by p omits salient and important facts. My response to the case where one is hit on the head with a hammer and hence irrationally believes one has just been hit on the head with a hammer was that, in this case, the fact believed would not supply the kind of explanation of one’s belief required by condition **K**. This is a development of Rieber’s response, intended to make clearer exactly when a fact is ‘salient’ and ‘important’ enough to be indispensable for the relevant explanatory purposes.

I have already said that I think Rieber's response is correct in essentials: it is indeed the 'inadequacy' (to an outsider) of the explanation of BAp by p alone that prevents the Newton case being a case of knowledge (although, as we saw, is not clear that Rieber is entitled to use this response). But Neta (667-8) claims that the response *does not work* in more complex cases. His argument for this claim is intriguing, and requires a rather intricate discussion which will occupy the rest of this section.

Neta imagines a more complex version of the case of Newton's apple:

An apple falls on Newton's head. By an amazing coincidence, Newton has a brain lesion such that, whatever rate the apple falls, the impact of the apple will (together with the lesion) cause Newton to believe that objects fall at precisely this rate *and that he has a brain lesion*. As it happens, the apple falls at 32ft./sec.², and Newton walks away with the belief that objects fall at this rate and that he has a brain lesion.

He argues that in this case:

Newton's conjunctive belief ... is explained by the fact that objects fall at 32 ft./sec.² and that he has a brain lesion. And yet even here Newton doesn't know that objects fall at 32ft./sec.², and this is something he must know if he knows the truth of the conjunction. Rieber cannot protest that we have still left out a "salient and important" part of the explanation: we can pack any such part of the explanation into the content of the explained belief as well. We thereby construct a case in which Newton's belief is caused by the fact believed but doesn't count as knowledge.

(I assume that ‘caused’ in the final sentence should read ‘explained’.) It may or may not be true that, in *Rieber’s* sense, Newton’s conjunctive belief is ‘explained’ by the fact believed in this case. But it is not *well explained to an outsider* just by that fact, as **K** requires. Let $p = \textit{objects fall at } 32 \textit{ ft./sec.}^2$, and let $q = \textit{Newton has a brain lesion}$. To see why **K** is not satisfied here, note that if we just cited $p \& q$ in explaining Newton’s belief that $p \& q$ to an outsider, we would mislead our audience. We would mislead them as to the way Newton acquired the belief that $p \& q$. By not mentioning that anything extraordinary is going on, we would lead our outsider to suppose that Newton acquired the belief that q in an ordinary way; perhaps by being told by his doctor that his brain was damaged, or by having noticed characteristic impairments in his own cognitive performance. A good explanation to an outsider of Newton’s belief that $p \& q$ will not simply cite $p \& q$, but must also make clear that the connection between $p \& q$ and Newton’s belief that $p \& q$ is an extraordinary one.

But what of Neta’s claim that such further explanatory information can simply be ‘packed into’ the content of the explained belief to generate new, genuine, problem cases? Well, let’s see what happens when we attempt to ‘pack in’ the appropriate information. Let $r = \textit{Newton believes } p \& q \textit{ in an irrational way after his apple-bump, because of his brain lesion}$. By ‘packing’ r into the content of Newton’s belief, we arrive at a case where Newton, once the apple has landed on his head, comes to believe (for no epistemically respectable reason) that $p \& q \& r$. For Neta’s objection to hold water against my account, it will have

to be the case that Newton's belief $p \& q \& r$ is well explained to an outsider just by $p \& q \& r$.

There are two points to make about this. First, look closely at what Newton supposedly believes in this case: he believes $p \& q$ and *that his belief that $p \& q$ is irrational*. This combination is psychologically unstable, and it is questionable whether it is really possible to believe a conjunction of claims which create such an obvious tension. So it is at least not obvious that Neta's method can be used to construct a *realistic* counterexample to my proposal.

This line of thought might be resisted, though; perhaps to some it will not sound so surprising to say that people hold beliefs which are in tension in the way described. After all, people will often say things of the form: 'p, but I don't suppose I have any good reason to think that'. However, even if we allow that it is in some sense possible for Newton to believe $p \& q \& r$, a second point prevents us from supposing that $p \& q \& r$ would then be, by itself, a good explanation to an outsider of his having that belief.

This is that the explanation would again be misleading: it does not make it clear that the third conjunct (r) is believed irrationally after the apple-bump because of the brain lesion. This must be mentioned, else we give the impression that Newton's belief in r is acquired normally (e.g. through the testimony of a rational and reliable informant), regardless of how the other beliefs are acquired. This point obviously iterates if we simply continue 'packing' new propositions like *r is believed irrationally* into the content of Newton's belief.

What if Newton believes $p \& q$ as before, but also believes r for normal reasons, e.g. on the basis of the testimony of a reliable informant? Then, we can say, he believes $p \& q \& r$ and there seems to be no reason to deny that $p \& q \& r$ is a good explanation of this belief. But he surely does not know p or q in this situation, although it seems we are forced to say he knows $p \& q \& r$, which (on any reasonable view) implies that he knows p and knows q . The earlier thought, that the explanation leaves out the irregularity of Newton's reasons for believing r , no longer applies, for those reasons are no longer irregular.

The explanation is nonetheless still misleading as to the relationship between explanandum and explanans. When presented with it, the outsider will not get a clear idea as to why Newton believes r . All of p , q , and r are needed to just to explain why Newton believes p and q . The explainer then intends r to be used *again*, quite separately, to explain (together with some background assumptions about likely ways for Newton to have learned that r) why Newton believes r . But the outsider would not have reason to see things that way, when presented with the simple explanation $p \& q \& r$ of Newton's belief. This simple explanation does not make it clear that r is being 'used twice' in this way. So the explanation is inadequate in this context.

To avoid this inadequacy, the explanation of Newton's belief will have to separate out the explanations for his believing these two parts of the proposition, and state that he believes $p \& q$ because $p \& q \& r$, and that he believes r because r . Since the simple explanation $p \& q \& r$ does not do this, it will not be a good

explanation to an outsider, condition **K** is not met in this case, and hence it is not a counterexample to my proposal.

The structure of an explanation matters, then. An explanation is not just a list of propositions, inadequacies in which can always be resolved by ‘packing’ more propositions onto the list. Neta’s objection cannot be pressed in this case without assuming a rather simplistic conception of what an explanation is.

There is one more possibility to be considered in our attempts to construct a counterexample along Neta’s lines. What if we try using self-referential propositions to correct for the problems with the previous attempt at a counterexample? Let $r' = \textit{Newton believes } p \& q \& r' \textit{ in an irrational way after his apple-bump, because of his brain lesion}$. Could a Neta-style case be contrived where Newton believes $p \& q \& r'$, and this belief is well explained to an outsider simply by citing $p \& q \& r'$? The problem described above will not arise: the idea here is that p and q and *part* of r' (the part that refers to p and q) are used to explain why Newton believes p and q , then the rest of r' (the part that refers to r' itself) is used to explain why he believes r' .

However, even if we assume for the sake of argument that sense can be made of the kind of self-referential belief required here, the explanation will still be inadequate. The outsider receiving $p \& q \& r'$ as an explanation of Newton’s belief that $p \& q \& r'$ is not given any clues as to the structure of the intended explanation, and so she could reasonably suppose that Newton’s belief in each conjunct is supposed to be explained by the corresponding conjunct of the explanans, which is not the case.

Given that several attempts to press Neta's objection have now failed, despite the construction of various *recherché* cases, I shall assume that this is not a promising line to pursue any further.

5 Other Theories and Intuitions

i) The Causal Theory

If some explanationist account of knowledge is correct, that reveals why a causal account is initially attractive for some kinds of knowledge (see Alvin Goldman 1967). For many, though not all, explanations are causal; and many, though not all, causal links are explanatory. The fact that the causal theory generates conditions on knowledge which are *prima facie* plausible but on reflection neither necessary nor sufficient suggests that some other concept, with which causation is loosely connected, could be used to generate a more promising account of knowledge. Explanation is an obvious candidate.

Roughly speaking, a causal theorist (see Goldman 1967) maintains that A knows that *p* just in case the fact that *p* causes A to believe *p*. By concentrating on explanation, we can see exactly why this sort of account goes wrong. Because some causal links are non-explanatory, a causal account of knowledge generates conditions which are insufficient. Such an account, as is well known, fails to make a principled distinction between the 'normal' and 'deviant' causal chains that can hold between our beliefs and the world. (This is why causal accounts are susceptible to Gettier-type counterexamples.) The explanationist can draw the

needed distinction by saying that ‘deviant’ causal chains are those where citing the cause would not provide a good explanation of the effect to an outsider.

And because some explanations are non-causal, the conditions generated by a causal account will also be non-necessary. It will have no natural way of dealing with (e.g.) knowledge of general truths such as *All emeralds are green*. For it is hard to see how such facts should be said to relate causally to our corresponding general beliefs. Note, however, that there is no reason to doubt that they can explain them. (I shall discuss this point further in §6 below.) Knowledge of future events is also prima facie problematic for causal accounts, as future events cannot cause present beliefs. But again, there is no obvious objection to the claim that facts about the future explain present beliefs. They can certainly explain present actions: I am tidying my flat today because my brother is coming to visit tomorrow. I also find it quite natural to say that I believe he is coming to visit tomorrow because he is coming to visit tomorrow.

It might be objected that what *really* explains my tidying my flat is the fact that I *believe* that my brother will visit tomorrow (a fact about the present). And similarly, what really explains my belief that he will visit tomorrow is that he has told me he will (a fact about the past). But what force would such an objection have? Bear in mind that all I need, for the current point to go through, is the claim that in certain contexts, the fact that my brother will visit tomorrow can constitute *a* good explanation of my tidying, and that similarly the fact that he will visit tomorrow can constitute *a* good explanation of my belief to that effect. I do not claim that these are *the* good explanations, so we need not worry if the

objector is just pointing out that the fact that I believe my brother will visit tomorrow *also* constitutes an adequate explanation of my tidying behaviour (or even the point that in most contexts it may be a fuller or better explanation). On the other hand, if there is meant to be more to the objection than this, I am not sure what it might be. If it is being claimed that the fact that my brother will visit tomorrow, although true, is never *a* good explanation of my tidying my flat, and that the fact that he will visit tomorrow is never a good explanation of my belief that he will, then I don't find the claim at all plausible.

ii) Nozick

It is quite plausible in general that if q is a good explanation of r then $\neg q \square \rightarrow \neg r$ (where ' $\square \rightarrow$ ' represents the subjunctive conditional). So, often, if p is a good explanation (to an outsider or anyone else) of BAp , then $\neg p \square \rightarrow \neg BAp$. This counterfactual is one of Nozick's conditions on knowledge. Nozick's proposal (in outline) is that KAp just in case:

1 p is true

2 BAp

3 $\neg p \square \rightarrow \neg BAp$

and 4 $p \square \rightarrow BAp$

Nozick's decision to include $p \square \rightarrow BAp$ as a fourth condition stems from his noticing that his conditions 1-3 are in some cases inadequate to guarantee knowledge. I shall suggest in this subsection that the source of Nozick's concern

is the feeling that in this kind of case, although there is some connection between p and BAp (which is why 3 obtains), there is something not quite right about that connection. This feeling, I shall say, is correct: in such cases, although there is some explanatory link between p and BAp , p fails to be a good explanation of BAp to an outsider, because there is something unusual about this link. Such cases must be ruled out by a successful account of knowledge. Nozick tries to rule out such cases by demanding $p \square \rightarrow BAp$, but I shall argue that this rules out too many kinds of unusualness in the connection between p and BAp , and thus generates familiar difficulties for Nozick. A major source of counterexamples to Nozick's account is the fact that, although 1-3 are not sufficient for p 's being a good explanation of BAp to an outsider, 4 is not necessary. This makes Nozick's account too strong. (Another problem is that there are a few cases where 1-4 hold but we do not have knowledge, so that the account is also too weak.) Nonetheless, Nozick's conditions are comprehensible from our perspective: 1-3 are, roughly, points of overlap with my proposal,^{vi} and 4 is a natural, if ultimately misguided, attempt to capture a further condition on explanatoriness.

To see that 4 is non-necessary, consider the case of the bystander who sees someone robbing a bank, and is able to identify him as Jesse James when his mask slips and she catches a glimpse of his face (Nozick 1981: 193). The case is difficult for Nozick, because his condition 4 ($p \square \rightarrow BAp$) is not met. In most close worlds where James robbed the bank, his mask did not slip and so the bystander formed no belief about who the culprit was. Yet we still want to say the bystander knows that James robbed the bank. (Nozick tries to resolve this

problem by appealing to the idea of a method of belief-formation, and specifying that all the relevant nearby worlds are ones where the same method is employed.)

I suggest that the case is best understood as akin to that of the wartime diary: the fact that James's mask slips is an unusual *feature* of the connection between the fact that James robbed the bank and the bystander's belief that he did, but it will not need to be mentioned in explaining her belief to an outsider, since *the connection itself* (visual perception in favourable conditions) is entirely normal. Hence we are inclined to say that the fact that James robbed the bank supplies a good explanation to an outsider of the bystander's belief that James robbed the bank. And therefore we feel that the bystander knows this. (If it had been the connection itself that was unusual – if, for instance, James's presence had triggered the projection of a hologramatic image of James, which the bystander then saw – we would not want to say that she knew James was there.)

To motivate 4, Nozick considers various cases where 1-3 are satisfied but we do not want to ascribe knowledge, because there is a 'deviant' link between the subject's belief and the fact believed. Amongst them is a case where a brain in a vat is given the belief that it is such by the vat-operator (see 1981: 175). All of 1-3 are satisfied here, yet the subject does not know it is a brain in a vat. 4 is supposed to help rule out such cases (although one might question whether it actually does so in all the relevant cases, as we shall see in a moment).

Unfortunately for Nozick, the condition $p \Box \rightarrow BAp$ rules out *too many kinds* of unusualness in the explanatory link between p and BAp , and this is what creates the problem in the Jesse James case. The condition $p \Box \rightarrow BAp$ rules out

cases where there is a normal link with unusual features, as well as cases where the link itself is unusual. The unusualness of the mask's slipping is only an unusual feature of the link between p and BAp ; the link itself is normal. Thus the unusualness of the mask's slipping does not interfere with the ascription of knowledge, because it does not interfere with our explaining BAp to an outsider by p alone. So condition 4, which rules out this kind of unusualness, is too strong. It forces Nozick to throw out the baby of glimpse-based knowledge with the bathwater of well informed vat-brains.

What Nozick needs to add to 1-3 is a condition which will rule out those cases where $\neg p \square \rightarrow \neg BAp$ holds simply because p is *part* of a good explanation to an outsider of BAp , although it is not such an explanation by itself (such as cases where there is a deviant causal link between p and BAp). To my knowledge, there is no better condition for this purpose than **K**.

It can furthermore be shown that 1-4 are not only too strong, in ruling out glimpse-based knowledge, but also too weak, in failing to rule out some things we *do* want to rule out. These are cases where there is an unusual explanatory connection between p and BAp such that the unusualness would have to be mentioned in explaining BAp to an outsider, but where the truth-values of p and BAp nonetheless covary across a range of close possible worlds, so that Nozick's conditions 3 and 4 are met. This might occur if I were a brain in a vat who formed a true belief that p on the basis of evidence fed into my brain by my benevolent vat-operator, who likes to give me true beliefs (and who has arranged for others to continue doing so should she be unable to). For then it seems that in

all close $\neg p$ -worlds, she, being benevolent, will make sure I do not have a belief that p , whereas in all close p worlds, she will make sure I have a belief that p . Here again, the explanatory constraint supplied by condition **K** seems to succeed where Nozick's conditions fail: this is not a case of knowledge because one would have to mention the strange link between p and BAp in explaining BAp to an outsider.

iii) Other Views

The reliabilist notion of 'belief acquired by a reliable method' is, in extension at least, fairly close to the notion of belief acquired in the sort of way an outsider would expect. Unless one is given a reason to suspect that something unusual is going on, one expects people to use reliable methods to form their beliefs. But the two sometimes differ in extension, because some reliable methods are unexpected (such as those employed by a brain in a vat who is reliably given true beliefs about its surroundings by a benevolent vat-operator). Where this happens, it is reliabilism that seems to be mistaken. (The vat-brain does not *know* what its surroundings are like, although it is acquiring true beliefs on that subject by a reliable method.)

Armstrong (1973) has suggested that A's knowing that p is a matter of the existence of a law-like connection between A's belief that p and the state of affairs which makes p true. If **K** is a genuinely necessary and sufficient condition for knowledge, then, given the influence of the covering-law model of

explanation, the emergence of an account like Armstrong's should not surprise us.

Let me briefly mention two other points of interest here. (These points could stand further development, but it would take me too far afield to pursue them in detail here.) Firstly, note that first-person beliefs about the sense data one is currently and consciously experiencing are plausible candidates for beliefs which can be well explained under any circumstances by citing the fact believed. Surely it will never be misleading, even when talking with an outsider, to explain my belief that there seems to me to be an orange patch in front of me just by saying that there seems to me to be an orange patch in front of me. The explanatory connection between fact and belief is very tight, and very obvious, in this case: there is no room for interference from the factors which usually cause trouble for explanatory links between facts and beliefs (such as misleading evidence). It is no coincidence, I think, that knowledge of one's present sense-experiences has sometimes been taken to be a most certain kind of knowledge.

Secondly, many philosophers think there is something troubling about knowledge of abstract objects. This is understandable when we consider that such objects could not interact causally with our beliefs, and that the concepts of causation and explanation have been (indeed, still are) often conflated and confused. This conflation makes it appear that facts about abstract objects could not possibly explain our beliefs about them. Hence, to anyone implicitly construing knowledge as requiring the satisfaction of condition **K**, knowledge of abstract objects would naturally appear problematic. (Whether or not there is a

genuine problem here is something I shall postpone until the final section of this paper.)

iv) Divided Intuitions

The proposed account of knowledge enables us to make sense of the fact that sometimes, when we are considering whether A knows that p, our intuitions are divided. Rieber has suggested that we can best understand a certain pattern of attitudes to scepticism by understanding particular aspects of the impact context can have upon explanation (1998, 195-7).^{vii} I agree, though I think we can do more to make the point perspicuous once condition **K** is in place.

The point rests upon the view that explanation is contrastive; that is, the view that when we explain why p, we are actually explaining why p rather than q for some foil q. Where a foil is not explicitly mentioned, an implicit foil is determined by contextual factors. (For a clear discussion of contrastive explanation see Lipton 1990.)

We can develop Rieber's own example (originally from Dretske 1970) to show how this view about explanation impacts upon sceptical debate in the light of condition **K**. Suppose Jane is looking at a zebra in a cage at the zoo. First we can note, with Rieber, that *This animal is a zebra* is a good answer to an outsider who asks 'Why does Jane believe that this animal is a zebra rather than a tiger?' but not to one who asks 'Why does Jane believe that this animal is a zebra rather than a cleverly disguised mule?'. To see how this helps us understand scepticism, note that in most contexts, we are only interested in how Jane forms beliefs when

confronted with animals that look the way they're supposed to look. So we are only interested in questions like the first. This sort of question is well-answered, even if an outsider asks it, by *This animal is a zebra*. We are therefore happy to say Jane knows the animal is a zebra. In some sceptical contexts, however, we interest ourselves in how Jane would respond to animals which have been cleverly disguised. We do not feel that *This animal is a zebra* is a good answer to an outsider who asks why she believes it is a zebra rather than a cleverly disguised mule. This answer is misleading to the outsider because it suggests that Jane can tell zebras from cleverly disguised mules, which she cannot. And so we no longer want to say Jane knows the animal is a zebra.

There doesn't seem to be a single right answer to the question of whether she knows it is a zebra or not. What we want to say is that 'in an everyday sense' she knows it, but she does not know it 'in the sceptic's sense'. And we can now spell out what is meant by this in terms of the two different types of why-questions which can be considered when asking whether condition **K** is met. (Whether or not we ought to endorse the pretheoretic claim that the word 'know' has different senses in the two contexts is a question that lies beyond the scope of this paper. It may be preferable to say that 'know' has a fixed sense, albeit one which builds in a contextual parameter; a discussion of this kind of issue can be found in Hawthorne 2004, especially chapter 4).

What is important for our purposes is that we want to say *Jane knows* when, and only when, we want to say *the fact that the animal is a zebra would be a good explanation to an outsider of her belief that it is a zebra*. Like all

accounts in this vein, my account of knowledge does not resolve our divided intuitions, or generate any kind of reply to the sceptic. What this type of account tries to do is help *explain* those divided intuitions, and help us *understand* scepticism.

Note that the proposal allows that simplistic epistemic closure claims (of the kind rejected by e.g. Dretske 1970, Nozick 1981 and Rieber 1998) may fail. It might be, for instance, that we ordinarily know that we have hands (= p) although we cannot know that we are not brains in vats (= q), and this despite the fact that we know that p entails q. For it is not the case that whenever r is a good explanation of BAr, and s is an obvious logical consequence of r, s is a good explanation of BAs. In the case of the two beliefs p and q just mentioned, the very nature of the shift in explanandum might be taken as forcing us to consider everyday standards of explanation, such as are usually considered correct for p, inappropriate for q.

6 Inductive Knowledge and Deductive Knowledge

The foregoing reflection might lead one to consider the following objection to the claim that **K** is necessary and sufficient for knowledge. Suppose A believes that p because she knows that q and either infers to p as the best explanation of q, or else deduces from q that p. In this sort of situation, even supposing that q is a good explanation to an outsider of BAq, we cannot assume that p is a good explanation to an outsider of BAp. Obvious candidate explanations for BAp include BAq, q, and either BA(*p is a good explanation of q*) or BA(*p is a*

consequence of q), but not *p*. But even if we don't allow that knowledge transmits across *all* good inductive or deductive inferences, it surely does in *some* cases. How can my proposal accommodate this?

Let us take the two cases of abduction and deduction separately. As regards the former, the solution to the problem is relatively straightforward. We simply note that in genuinely knowledge-producing cases of inference to the best explanation, where *p* explains *q*, *q* explains *BAq*, and *BAq* explains *BAp*, and all these explanatory links are standard, then one can collapse the explanatory chain and simply say that *p* explains *BAp*, even when talking to an outsider.

For example, if I believe that all emeralds are green on the grounds of an inference from known instances of green emeralds, it seems reasonable to say that the fact that all emeralds are green is what explains my belief that they are. For it is what explains the fact that all the emeralds I have seen so far have been green, and this is what explains my general belief. The central steps in the explanatory link between fact and belief are exactly what anyone with no special knowledge of my situation would expect them to be, so they can be safely omitted without rendering the explanation inadequate to such a person.

We do not need to worry about cases where *p* is *not* really a good explanation of the known fact(s) *q* from which *A* infers, because in these cases inferring *p* as the best explanation of *q* surely does not generate knowledge that *p*. But what of cases where *p* is a good explanation of *q* to the subject *A*, but not to an outsider? This can happen when *A* has information which is relevant to her understanding of the relationship between *p* and *q*, but which is not common

knowledge. These cases need not trouble us either, however. Given that A reasons to p as the best explanation of q, which the outsider would not do, there are two possibilities: either A is *better* informed than the outsider, or A is *misinformed*. In situations where he is better informed, it is acceptable to explain BAp to an outsider by p alone. For the outsider will assume some rational process has lead to A's believing p, inference to the best explanation being one possibility. And the outsider will assume that, if A has inferred to p as the best explanation of q, that is because p *really is* a good explanation of q (and A is aware of this). For unless one has specific information about a person, one assumes that that person will make good abductive inferences. So the explanatory chain mentioned above ($p \rightarrow q \rightarrow \text{BA}q \rightarrow \text{BA}p$) is once again collapsible for the outsider, since in particular the collapsed explanation will not mislead the outsider by suggesting that, if A has arrived at the belief that p by abductive means, then the abductive inference she used was a good one. Cases where A takes p to be a good explanation of q because he is *misinformed* need not detain us, since they will not be cases of knowledge. (At best, they will be cases of accidentally justified true belief).

Deductive cases cannot be addressed in quite this way. We cannot assume that just because p is a consequence of q, p would be a good explanation of q to an outsider (or indeed, to anyone). So the first step in the chain by which p explains BAp in abductive cases cannot be assumed to hold in deductive cases. Nor can we simply stipulate that all propositions deduced from known propositions are known. For it seems that knowledge may sometimes fail to be

transmitted in this way. For instance, as noted in §5 above, it has seemed plausible to some philosophers that we know we have hands but do not know we are not brains in vats, although the latter is a known consequence of the former.

Nonetheless, it does seem that those cases where we want to say knowledge *is* transmitted across a deductive inference are precisely those where we are prepared to say that the resulting belief BAp would be well explained to an outsider by p. We can best see this by considering a concrete case.

Suppose Brian, a competent logical reasoner, knows that if Neil comes to his party he will get drunk, and knows that Neil is coming to his party. He deduces that Neil will get drunk. Is the fact that Neil will get drunk a good explanation to an outsider of Brian's belief to that effect? Well, first note it is only because Neil *will* get drunk that the two premises from which Brian reasons are true together; that is to say, it is only because Neil will get drunk that Brian's argument is sound. What kind of 'because' are we dealing with here? One, I would suggest, which signals a familiar kind of explanation. For many explanations state the fulfilment of an obvious necessary (or almost-necessary) condition on the truth of the explanandum. For instance, I can explain why I arrived on time today by saying that, for once, nothing held me up. (But note that not all such conditions can be used in this way: more on this in a moment.)

Thus we can explain the soundness of Brian's argument by citing the fact that Neil will get drunk. In a similar vein, after noting that competent reasoners like Brian will (generally) only use sound arguments, we can explain the fact that Brian reasons as he does by citing the soundness of Brian's argument. Again, we

are explaining the truth of the target proposition by stating that an obvious (almost-) necessary condition for it is met.

Finally, we note that, uncontroversially, Brian's reasoning as he did is what explains his belief that Neil will get drunk. So we now know that the fact that Neil will get drunk explains why Brian's argument is sound, which explains Brian's reasoning as he did, which in turn explains why Brian believes that Neil will get drunk. And the central stages in this explanatory link are entirely standard: people very often reason according to *modus ponens* in this sort of way. So these central stages can be omitted without making the explanation misleading, even when talking to an outsider. In other words, the fact that Neil will get drunk is, by itself, a good explanation to an outsider of Brian's belief that he will. And this is why we want to say Brian knows the Neil will get drunk. Of course, if Brian were an incompetent logician who drew conclusions randomly from his premises, or if he were using an unsound argument, I could not argue as I have done. But then we would not want to say that Brian knew Neil would get drunk.

A qualification is needed here, however. Note that explaining a fact by citing an obvious necessary condition only seems to work where it is, or at least could be, less obvious that the explanans obtains than that the explanandum does. In our example, it is less obvious (to Brian at least) that Neil will get drunk than that the relevant instance of *modus ponens* is sound. (That is why Brian uses this instance of *modus ponens* to deduce that Neil will get drunk.) Similarly, my not

having been held up today must be less obvious than the fact that I got here on time, if we are to explain the latter by the former.

Compare this with a typical case where, it is argued, warrant does not transmit across deductive inference. Michelle believes she has two hands and that brains in vats do not have hands, and she deduces that she is not a brain in a vat. What goes wrong here is that can be no more obvious (to Michelle or anyone else) that she has two hands than that she is not a brain in a vat. And since knowing she has two hands is necessary for knowing that her argument is sound, it can be no more obvious that her argument is sound than that she is not a brain in a vat. This means we cannot satisfactorily explain to an outsider why she reasons as she does by citing the truth of her conclusion, as we could in Brian's case.

This analysis of the situation is, I think, confirmed by the fact that philosophers who argue that warrant does transmit across this kind of entailment (e.g. Moore 1959) assume that facts concerning our having hands *are* more obvious than facts concerning our not being brains in vats. Their opponents resist the transmission claim by arguing that this is not so (see e.g. Davies 2000, 400-1).

7 Conclusions

A significant advantage of my proposal, which I have not yet mentioned, is that it may help to render unmysterious the existence of logical and mathematical knowledge. There is reasonably widespread agreement that at least some necessary truths can serve as explanations: for discussions of mathematical

explanation, see e.g. Steiner 1978, Resnik and Kushner 1987, and Sandborg 1998. Although there exists as yet no agreed account of the nature even of mathematical explanation, if it is so much as possible for the account sketched in this paper to be helpful when we try to understand our knowledge of necessary truths, that represents a significant advantage over causal accounts (which seem straightforwardly inapplicable to mathematical and logical knowledge; see Goldman 1967: 362) and Nozick's account (clause 3 of which is, on the standard semantics for the counterfactual, trivially satisfied whenever one believes a necessary truth). Of course, much work needs to be done if we are to show *how* a logical or mathematical truth could explain our having a corresponding belief. One might, for instance, have serious doubts about the plausibility of the idea that we can explain the contingent, down-to-earth fact that I believe that $7+5=12$ by appeal to a necessary proposition which deals in abstract objects (or at least appears to).

Worries of this sort may be surmountable, however. In Jenkins (forthcoming), I suggest that arithmetical truths are known through an examination of arithmetical concepts which are epistemically grounded in experience. The concepts in question are taken to be reliable epistemic guides to arithmetical truth, because they are sensitive to the way our sensory input goes, and hence sensitive to the way the external world is. If something like this is correct, we can posit a three-stage explanation between arithmetical fact and arithmetical belief. First, the arithmetical fact explains why, in the relevant respects, our sensory input goes the way it does. Second, the fact that our sensory

input goes the way it does explains why we possess the arithmetical concepts we do. And third, our possession of those concepts explains our belief in conceptually true propositions involving those concepts. Since all the stages in this link are (on the proposal under consideration) perfectly rational and normal, it would be permissible to collapse the explanation and cite the arithmetical fact alone as an explanation of the belief, even when talking to an outsider.

In this paper I have suggested that a belief counts as knowledge iff the subject's having that belief can be explained, to someone with no special knowledge of the subject's situation, just by citing the fact believed. I have not presented this suggestion as an analysis of the concept of knowledge, but as a way of dramatizing the notion of 'normality' which is so closely bound up with that concept, and which (it seems) any interesting account of what knowledge is must somehow help elucidate.

I said at the beginning of the paper that the aim of the condition on knowledge to be offered here was to help us better understand the rôle that the concept of knowledge plays in our lives (the project begun in Craig 1990). Let me conclude with a few remarks on this.

Traditionally, epistemologists sought a simple and reductive answer to the question of what knowledge is. But there is no reason to suppose there will be a simple reductive answer to this question; knowledge might simply be 'normal' true belief, and what counts as 'normal' might simply be a function of the (complex and diffuse) requirements that we generally want our beliefs to satisfy.

However, this doesn't mean we can't say anything helpful about the notion of normality in play. One thing we can do is ask how the normality concept in play here relates to other concepts in regular use. What I have attempted to do is to relate the normality concept to the familiar (if philosophically intriguing) concept of explanation. If it is true that 'normal' true beliefs are just those of which **K** holds, then we may be able to learn more about what we tend to count as a 'normal' belief by considering under what circumstances instances of **K** are true.

To do this, we need to explore as far as possible our attitudes towards explanation, explanatory value and the impact thereon of contextual parameters. And such work is, of course, already being done. My aim here is simply to point out that clues as to what knowledge is and what sort of rôle the concept of knowledge plays in our lives may be uncovered by applying the results of this investigation.^{viii}

References

- Armstrong, D. 1973. *Belief, Truth and Knowledge*. Cambridge: Cambridge University Press.
- Craig, E. 1990. *Knowledge and the State of Nature*. Oxford: Clarendon Press.
- Davies, M. 2000. "Externalism and Armchair Knowledge." In *New Essays on the A Priori*, ed. P. Boghossian and C. Peacocke, Oxford: Clarendon Press, 384-414.
- Dretske, F. 1970. "Epistemic Operators." *Journal of Philosophy* 67 (1970): 1007-1023.
- Garfinkel, A. 1981. *Forms of Explanation*. New Haven, Conn.: Yale University Press.
- Gettier, E. 1963. "Is Justified True Belief Knowledge?" *Analysis* 23: 121-23.
- Goldman, Alan. 1984. "An Explanatory Analysis of Knowledge." *American Philosophical Quarterly* 21: 101-108.
- 1988. *Empirical Knowledge*. Berkeley: University of California Press.
- Goldman, Alvin. 1967. "A Causal Theory of Knowing." *Journal of Philosophy* 64: 355-72.
- Hawthorne, J. 2004. *Knowledge and Lotteries*. Oxford: Clarendon Press.
- Jenkins, C. Forthcoming. "Knowledge of Arithmetic." *British Journal for the Philosophy of Science*. Available online at: <http://www.st-andrews.ac.uk/~csj6/>.
- Lipton, P. 1990. "Contrastive Explanation." In *Explanation and its*

- Limits, ed. D. Knowles, Cambridge: Cambridge University Press. Reprinted in *Explanation* (1993), ed. D.-H. Ruben Oxford: Oxford University Press.
- Moore, G.E. 1959. "Proof of an external world". In his *Philosophical Papers*, London: Allen and Unwin.
- Neta, R. 2002. "S Knows That P." *Noûs* 36: 663-81.
- Nozick, R. 1981. *Philosophical Explanations*. Oxford: Clarendon Press.
- Resnik, M. and Kushner, D. 1987. "Explanation, Independence and Realism in Mathematics." *British Journal for the Philosophy of Science* 38: 141-58.
- Rieber, S. 1998. "Skepticism and Contrastive Explanation." *Noûs* 32: 189-204.
- Sandborg, D. 1998. "Mathematical Explanation and the Theory of Why-Questions". *British Journal for the Philosophy of Science* 49: 603-24.
- Steiner, M. 1978. "Mathematical Explanation." *Philosophical Studies* 34: 135-51.
- Van Fraassen, B. 1980. *The Scientific Image*. Oxford: Clarendon Press.
- Williamson, T. 2000. *Knowledge And Its Limits*. Oxford: Oxford University Press.

Notes

ⁱ Besides Craig, notably Williamson (2000: 2-5).

ⁱⁱ All further occurrences of ‘Goldman’ refer to Alan Goldman unless otherwise specified.

ⁱⁱⁱ Note that I assume it is not necessary, for p to be a good explanation of BAp in a context C , that p *actually be cited* in C as an explanation of BAp. It is sufficient that if it were so cited it would be a good explanation.

^{iv} This seems to me a spurious reason for rejecting the phrase ‘good explanation’, and so I use it freely in my account. Even if explanatory goodness comes by degrees, so that (e.g.) some explanations are ‘fairly good’ and some are ‘not very good’, there seems to be a (perhaps vague, but at least reasonably well-understood) boundary above which an explanation is ‘good’ simpliciter and below which it is not.

^v Thanks to an anonymous referee for bringing the general point, and this particular case, to my attention.

^{vi} Condition 3, however, creates some difficulties which my account can avoid. For instance, the case of the concerned Granny (see Nozick 1981, 179) forces Nozick to amend condition 3 so as to take account of the ‘method’ by which a belief is acquired. We can accommodate the concerned Granny by noting that 3, although a symptom of the satisfaction of condition **K**, is not a necessary condition for **K** in all cases.

^{vii} Neta also provides a helpful discussion of how contextualism about explanation can help us understand various puzzles about knowledge, which he lists on p. 664.

^{viii} My thanks to Simon Betts, Simon Blackburn and Brian King for helpful discussions on this topic. I am also grateful to those who participated in a Graduate Seminar in Cambridge where an early draft of this paper was read in April 2002, and to two anonymous referees for this journal who provided helpful and constructive comments.