

# UC San Diego

## UC San Diego Previously Published Works

### Title

Williamson on knowledge and psychological explanation

### Permalink

<https://escholarship.org/uc/item/36m476jv>

### Journal

Philosophical Studies, 116(1)

### ISSN

0031-8116

### Authors

Magnus, P D  
Cohen, Jonathan

### Publication Date

2003-10-01

Peer reviewed

# Williamson on Knowledge and Psychological Explanation\*

P. D. Magnus<sup>†</sup> and Jonathan Cohen<sup>‡</sup>

## Abstract

According to many philosophers, psychological explanation can legitimately be given in terms of belief and desire, but not in terms of knowledge. To explain why someone does what they do (so the common wisdom holds) you can appeal to what they think or what they want, but not what they know. Timothy Williamson has recently argued against this view. Knowledge, Williamson insists, plays an essential role in ordinary psychological explanation. Williamson's argument works on two fronts. First, he argues against the claim that, unlike knowledge, belief is "composite" (representable as a conjunction of a narrow and a broad condition). Belief's failure to be composite, Williamson thinks, undermines the usual motivations for psychological explanation in terms of belief rather than knowledge. Unfortunately, we claim, the motivations Williamson argues against do not depend on the claim that belief is composite, so what he says leaves the case for a psychology of belief unscathed. Second, Williamson argues that knowledge can sometimes provide a better explanation of action than belief can. We argue that, in the cases considered, explanations that cite beliefs (but not knowledge) are no less successful than explanations that cite knowledge. Thus, we conclude that Williamson's arguments fail both coming and going: they fail to undermine a psychology of belief, and they fail to motivate a psychology of knowledge.

Is knowledge the sort of state that should figure in explanations of action? According to many philosophers who think that folk psychological explanations are important and generally successful, the answer is no. Rather, they think, folk psychological explanation should advert to beliefs and desires, but not to knowledge. Recently, Timothy Williamson [Williamson, 2000] has argued that this view is erroneous — that knowledge does play a role in ordinary psychological explanation. Williamson's arguments are aimed to deliver a one-two punch to opponents. One argument, if successful, would defeat common reasons for thinking that belief has more explanatory power than knowledge. Another

---

\* *Revision* : 1.10

<sup>†</sup>Department of Philosophy, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0119, pmagnus@fecundity.com

<sup>‡</sup>Department of Philosophy, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0119, joncohen@aardvark.ucsd.edu

would turn the opposing view on its ear by showing that knowledge has more explanatory power than belief. In §1, we discuss Williamson’s argument against the view that only belief can be causally efficacious in producing action. In §2, we address Williamson’s argument for the conclusion that knowledge explains action better than belief does.

## 1 Explanation and Primeness

### 1.1 Background: Causal Explanation

The dispute over the role of knowledge in folk psychological explanation is best appreciated against the background of a general understanding of (folk) psychological explanation that many have found appealing in some form.<sup>1</sup> This picture has it that psychological explanation, like explanation in the special sciences generally, involves subsuming individuals under *ceteris paribus* causal generalizations. Suppose, reasonably, that causal generalizations subsume individuals in virtue of their causal powers. Then it follows that psychological generalizations must taxonomize individuals in terms of their causal powers. But, it is widely supposed, causal powers must be intrinsic, or at least have intrinsic grounds; as McGinn puts the the point,

... what happens at the causal nexus is local, proximate and intrinsic: the features of the cause that lead to the effect must be right there where the causal interaction takes place. Causation is the same with brains and minds as it is with billiard balls. Their effects depend on local properties of these entities. The causal powers of a state or property must be intrinsically grounded; they cannot depend essentially upon relations to what lies quite elsewhere ([McGinn, 1989], 133).

Putting this all together, we get the conclusion that the properties that figure in psychological explanation must be intrinsic, or at least have intrinsic grounds: for a property to figure in psychological explanation, it must be a property that supervenes on what is inside the heads of subjects.<sup>2</sup>

On its face, however, knowledge does not seem to be such a property. Knowledge entails truth — a subject cannot know that  $p$  if  $p$  is false. But, as they say, truth ain’t in the head: generally speaking (that is, excepting  $ps$  about what is happening in the heads of subjects), whether  $p$  is true fails to supervene on

---

<sup>1</sup>In attempting to situate Williamson’s discussion in the wider literature, we may have failed to capture his intentions exactly (he says relatively little to characterize the positions against which he is arguing). We believe that this speculation is warranted (even if risky) insofar as it clarifies what is at stake for Williamson.

<sup>2</sup>Fodor’s famous version of this argument is grounded in a “formality condition” — a condition that, in effect, recapitulates the assumptions above about causal powers being intrinsic. He concludes that “Since ... knowledge is involved with truth, and since truth is a semantic notion, it’s going to follow that there can’t be a psychology of *knowledge* (even if it is consonant with the formality condition to hope for a psychology of *belief*)” ([Fodor, 1980], 64).

what is happening in any subject’s head, but does “depend essentially upon relations to what lies quite elsewhere.” Given that knowledge entails truth, and that truth ain’t in the head, it follows that knowledge ain’t in the head either. Consequently, on the assumptions about psychological explanation rehearsed above, knowledge is not the sort of property that can figure in psychological explanation.

## 1.2 Externalism, Belief, and Knowledge

Unfortunately, things are not so simple as they might seem so far. For the foregoing arguments for the conclusion that knowledge can’t figure in psychological explanations also pose a *prima facie* threat to the viability of psychological explanation in terms of belief. For, on widely held externalist assumptions about mental content, whether you count as believing that  $p$  fails to supervene on what is in your head. However, many have thought that this threat could be answered by a divide and conquer strategy. In particular, the thought goes, we can think of belief as built up from a narrow factor, which supervenes on what is in the head, together with a broad factor, which fails to supervene on what is in the head.<sup>3</sup> Thus, on one influential version of this view (cf. [Fodor, 1975], 75, [Fodor, 1987], 17, [Schiffer, 1987], chapter 4), a subject  $S$  believes that  $p$  iff:

(NF)  $S$  stands in some computationally specified relation to a mental representation  $\phi$ , and

(BF)  $\phi$  means that (/has the content that)  $p$ .

As the labels are intended to suggest, (NF) is a narrow factor and (BF) is a broad factor: whether (NF) holds supervenes on what is inside  $S$ ’s head, while, if externalism is right, whether (BF) holds does not.

How can this understanding of belief as composed from narrow and broad factors rescue the idea that beliefs are causally potent? The idea is that, even if belief ain’t in the head, there is something in the head that will do for purposes of psychological explanation — viz., the mental representation  $\phi$  that figures in the foregoing account of belief. Since  $\phi$  is in the head,  $\phi$  can serve as a proxy for belief in causal explanations: thus, even if the belief that  $p$  can’t be causally efficacious on its own, it can be accorded a derivative sort of causal efficacy in terms of the causal efficacy of the (narrow) mental representation  $\phi$  that means that  $p$ . And since the causal efficacy of the latter is compatible with the considerations about the locality of causation bruited above, everything will work out fine.<sup>4</sup>

<sup>3</sup>In what follows we use standard jargon (including the terms ‘narrow’ and ‘environmental’) and some reasonably intuitive extensions thereof. For Williamson’s explicit definitions of these terms, along with ‘prime’, ‘composite’, ‘internally like’, and ‘externally like’, see p. 66.

<sup>4</sup>Notice also that, since the fundamental (non-derivative) causal work is done by a narrow factor, individuals who are narrowly alike but environmentally not alike (say, you and your twin on Twin Earth) can be subsumed by the same causal explanations. On this point, as well as a fuller development of the move rehearsed here, see [Kim, 1982], [Fodor, 1987], chapter 2, and [Jackson, 1996].

The lesson seems to be this. If a propositional attitude state is to figure in psychological explanations, it must be causally efficacious. But propositional attitude states can't be causally efficacious except in a derivative sense, wherein their causal efficacy is explained in terms of the causal efficacy of their narrow components (in order to respect locality requirements). Consequently, if a propositional attitude state is to figure in psychological explanations, it must be possible to isolate that state's narrow component; and this requires that such a state can be factored into narrow and broad components.

Suppose we accept this lesson, and thereby agree that a propositional attitude state cannot play a role in psychological explanation unless it can be factored into narrow and broad components. As we've seen, this lesson is compatible with taking belief to play a robust role in psychological explanation, insofar as there are available accounts that factor belief into narrow and broad components. Could we say the same thing about knowledge? Many epistemologists hold that knowledge can be factored in the relevant sense; however, for these epistemologists, the narrow component of knowledge is either belief itself, or belief together with some further narrow condition (e.g., justification).<sup>5</sup> But if this is right, then attributions of causal efficacy to knowledge work, if at all, by piggybacking on the causal efficacy of belief. And to say this is to admit that explanations in terms of knowledge are, at best, longhand recapitulations of explanations in terms of belief; moreover, they would provide recapitulations of only a limited subset of the explanations that a psychology of belief makes available. Consequently, on these assumptions, it is hard to see why we would want a psychology of knowledge, even if we could have one.<sup>6</sup>

### 1.3 Primeness

The argument just reviewed holds that knowledge can only serve in psychological explanation if it can be factored into narrow and broad components, and that if it can be so factored, then the way it factors makes it less useful for psychological explanation than belief. However, Williamson seems unmoved by this argument; indeed, he doubts that any considerations involving factorization can choose between a psychology of belief and a psychology of knowledge. For, he thinks, both belief and knowledge are "prime" rather than "composite" — he thinks neither can be represented as the conjunction of a narrow condition and an environmental condition. If he is right about this, then belief and knowledge are on all fours as far as the requirements of causal efficacy are concerned, so (contrary to the consensus) such considerations do not favor psychological explanations in terms of belief over psychological explanations in terms of knowledge. Williamson has a very general form of argument that he uses to establish that belief, knowledge, and many other mental states, are prime; we'll first sketch

---

<sup>5</sup>See [Plantinga, 1993] for a survey of accounts of knowledge that would underwrite this claim.

<sup>6</sup>As an example of this line of thought, Williamson cites Stich's remark that "what knowledge adds to belief is psychologically irrelevant" ([Stich, 1978], 574; quoted approvingly in [Kim, 1982], 188).

his argument schematically, and then turn to the instance of the argument that concerns belief.<sup>7</sup>

Williamson's general argument form begins by supposing for *reductio* that a mental state  $S$  is composite; if so, then  $S$  is the conjunction of a narrow condition  $N$  and a broad condition  $B$ . He hopes to generate a contradiction from the hypothetical assumption by finding three cases  $\alpha$ ,  $\beta$ , and  $\gamma$  such that

- (i)  $S$  obtains in  $\alpha$  and  $\beta$ ,
- (ii)  $\alpha$  and  $\gamma$  are internally alike,
- (iii)  $\beta$  and  $\gamma$  are externally alike, and
- (iv)  $S$  fails to obtain in  $\gamma$ .

(i) and (ii) entail that  $N$  obtains in  $\gamma$ , and (i) and (iii) entail that  $B$  obtains in  $\gamma$ , so  $N$  and  $B$  obtain in  $\gamma$ ; but, on the assumption that  $S$  is composite, this contradicts (iv).

Thus, a combination of cases  $\alpha$ ,  $\beta$ , and  $\gamma$  that meets (i)–(iv) is inconsistent with the supposition that  $S$  is composite. Now, Williamson chooses  $\alpha$  and  $\beta$  as two possible realizations of  $S$ ; on the current assumption that  $S$  is composite, this means that there is a narrow condition  $N_\alpha$  and a broad condition  $B_\alpha$  such that  $\alpha$  holds just in case  $(N_\alpha \& B_\alpha)$  holds, and a narrow condition  $N_\beta$  and a broad condition  $B_\beta$  such that  $\beta$  holds just in case  $(N_\beta \& B_\beta)$  holds. (He also chooses conditions so that  $N_\alpha$  and  $N_\beta$  are mutually exclusive and  $B_\alpha$  and  $B_\beta$  are mutually exclusive.) This will guarantee (i). But how does Williamson ensure that there is a case  $\gamma$  that satisfies (ii)–(iv)? On the assumption in place for the *reductio*,  $S$  is composite, so the internal components  $N_\alpha$  and  $N_\beta$  can vary independently of the external components  $B_\alpha$  and  $B_\beta$ ; therefore, we may construct  $\gamma$  such that  $(N_\alpha \& B_\beta)$  holds in  $\gamma$ .  $\gamma$  is internally like  $\alpha$  (as per (ii)) and externally like  $\beta$  (as per (iii)). But since neither  $(N_\alpha \& B_\alpha)$  nor  $(N_\beta \& B_\beta)$  holds in  $\gamma$ , neither of the two possible realizations of  $S$  holds in  $\gamma$ , so  $S$  will fail to hold in  $\gamma$ .

Williamson wants to urge that belief, like knowledge, is prime. Consider, then, the sort of view attributed to Fodor above, on which having the belief that  $p$  — say, believing that tigers growl — is a matter of meeting conditions (NF) and (BF) above. Williamson argues against this view by applying his general argument form to the case at hand:

Let  $\alpha$  be a case in which tigers inhabit the mountains while schmiglers (which appear just like tigers) inhabit the jungle; one remembers one's encounters with tigers in the mountains but totally forgets

---

<sup>7</sup>Williamson applies this general argument form to a large number of states: the state of seeing water (69–70), the state of hearing a certain individual Mary (70), the state of believing that a particular screen flickers (70–71), the state of believing that tigers growl (71–72), the state of knowing that tigers growl (72), and the state of knowing by testimony that a certain election was rigged (72). The argument form is indeed widely applicable and, we believe, very elegant.

one's encounters with schmigers in the jungle. One believes that tigers growl; since one has no recollection of schmigers, one does not believe that schmigers growl. Let  $\beta$  be a case in which tigers inhabit the jungle while schmigers inhabit the mountains; one remembers one's encounters with tigers in the jungle but totally forgets one's encounters with schmigers in the mountains. One believes that tigers growl; since one has no recollection of schmigers, one does not believe that schmigers growl. Now consider a case  $\gamma$  internally like  $\alpha$  and externally like  $\beta$ . In  $\gamma$ , tigers inhabit the jungle while schmigers inhabit the mountains; one remembers one's encounters with schmigers in the mountains but totally forgets one's encounters with tigers in the jungle. One believes that schmigers growl; since one has no recollection of tigers, one does not believe that tigers growl. Thus the condition that one believes that tigers growl is prime (71).

We believe this argument is ineffective against the sort of view we have discussed. To see why, notice that (NF), as stated, is a quantificational statement; it says that there is some mental representation  $\phi$  such that  $\phi$  stands in a certain relation to  $S$ . (BF), then, is not an independent conjunct tacked onto (NF), as it might appear to be on a first reading (we are not suggesting that Williamson is unaware of this — see below); rather, it adds a condition onto  $\phi$  — that very mental representation that makes (NF) true. To put the point in explicitly logical terms, the additional condition that (BF) imposes occurs within the scope of the quantifier introduced in (NF).

We believe that the apparent force of Williamson's argument overlooks this point: his argument depends essentially on constructing a case  $\gamma$  by varying independently narrow and broad factors that are taken from  $\alpha$  and  $\beta$ . If we follow the account of belief under consideration in taking these factors as bound by a single quantifier, Williamson's argument collapses. To see this, consider  $\alpha$ ,  $\beta$ , and  $\gamma$  again in the context of the (NF), (BF) account of believing that tigers growl. It does seem that  $\alpha$  is internally like  $\gamma$  — in both cases one stands in the same relevant computationally individuated relation to a (narrowly individuated) mental representation of the same type; to use a familiar metaphor that Williamson adopts for vividness, we may say that in both  $\alpha$  and  $\gamma$  one's belief box contains a token of the same (narrowly individuated) Mentalese expression type. Are  $\beta$  and  $\gamma$  externally alike? Williamson's argument suggests that we should answer this question affirmatively on the grounds that the tigers and schmigers are in the same place in  $\beta$  and  $\gamma$  (tigers in the jungle, schmigers in the mountains). But on the account of belief under consideration, that is not the way to decide whether  $\beta$  and  $\gamma$  are externally alike. The question we should be asking is one that cannot be articulated in terms of a broad factor alone, because (on the theories at issue) the broad factor and the narrow factor are bound by a single quantifier. We should be asking this question: Is the particular token in the belief box in  $\beta$  (appropriately) causally related to the very same thing to which the particular token in the belief box in  $\gamma$  is (appropriately) causally related? The answer to this question is, of course, negative: in  $\beta$  the

(unforgotten) token in the belief box is causally related to tigers, while in  $\gamma$  the (unforgotten) token in the belief box is causally related to schmigers. Thus,  $\beta$  and  $\gamma$  are *not* externally alike in the sense that matters to the kinds of theories at issue, so  $\gamma$  is not, as advertised, a counterexample to those theories.

We take the above considerations to show that Williamson's argument wins him, at best, a Pyrrhic victory. His argument shows that belief cannot be understood as the conjunction of a narrow factor and a broad factor, where that would mean that the factors can be satisfied or not satisfied independently of one another. However, this conclusion is not damaging to the sorts of accounts of belief, knowledge, and other states that philosophers have proposed, and that Williamson's terminology (e.g., the appeal to a belief box and expressions in a language of thought) suggest that he intends to address. These views are compatible with the claim that belief is prime in Williamson's sense. For, while these views require that belief can be factored into narrow and broad components, they do not require that the two factors can be satisfied or not satisfied independently of one another. The relevant sort of factorization, then, does not presuppose compositeness; consequently, Williamson's argument that belief is not composite leaves untouched accounts that require factorization.

Williamson is aware of the possibility of this response, but, oddly, he gives only a brief discussion of the matter at the end of the relevant chapter (89–92). There he admits that his argument is ineffective against accounts on which the narrow factor and the broad factor are held to be connected rather than capable of independent variation. However, he claims, this concession is of no help to his opponent in the context of the project of understanding psychological explanation unless the narrow factors enlisted in the analysis of belief are adequate to the needs of psychological explanation. For reasons we shall consider in what follows, Williamson doubts that they are.

Before we come to this, however, we should pause to note that the criticism we are making here is broadly in agreement with that in [Brueckner, 2002]. Brueckner considers accounts that factor knowledge into broad and narrow components that are causally related rather than independent. Williamson's insistence on the primeness of knowledge, he argues, is not damaging to such accounts. Our diagnosis is more general than Brueckner's in two significant respects: First, our criticism is that the primeness arguments fail if the broad and narrow components stand in *any* relation that precludes their independent variation — be it causal or otherwise. The constraint results from the logical form of such accounts, since (NF) and (BF) are within the scope of a single quantifier. Second, Brueckner is concerned exclusively with the application of the argument to the notion of knowledge. It is important to our argument here that the diagnosis should apply quite generally to all instances of Williamson's argument form, including its application to accounts of belief.<sup>8</sup>

---

<sup>8</sup>Thanks to Anthony Brueckner for helpful discussion of these connections.



## 2 The Explanatory Role of Knowledge

As we take it, Williamson’s argument has two parts: a ‘primeness’ part that is supposed to establish that knowledge is in no worse than belief as an explainer, and a probabilistic part that is supposed to establish that knowledge is in fact a better explainer. In the previous section, we found the ‘primeness’ argument unconvincing. What, then, of the probabilistic argument?

As we have seen, Williamson thinks that belief and knowledge are equally prime, and therefore that the usual line of argument does not motivate a conception of psychological explanation that appeals only to belief rather than knowledge. For all that, one might still think that psychology *could* mention knowledge but wouldn’t *need* to. On the contrary, Williamson argues, knowledge is sometimes *more* explanatorily useful than belief is. That a person knows something, he claims, may better explain her actions than merely that she believes it truly. Knowledge (so Williamson claims) is indispensable for psychological explanation in general, even though it does not need to be invoked exclusively or in every psychological explanation; that is, knowledge must be included alongside belief in the conceptual toolbox we use in explaining action.

### 2.1 Formal apparatus

Consider the general case of a subject who entertains an action that would be appropriate under some circumstance.<sup>9</sup> Let  $A$  be the condition that the subject performs the action, let  $p$  be the appropriate circumstances for  $A$  to obtain, let  $B$  be the condition that the subject believes  $p$ , and let  $K$  be the condition that the subject knows  $p$ . Define  $B'$  as the condition that the subject believes truly but does not know:

$$B' \stackrel{def}{=} B \ \& \ p \ \& \ \neg K \tag{1}$$

Williamson claims that for a significant range of cases  $K$  will be more highly correlated with  $A$  than is  $B'$ . Thus, we would be able to accurately predict  $A$  more effectively by advertent to knowledge than by advertent to true belief.<sup>10</sup>

Williamson wants to argue that the correlation between  $A$  and  $K$  is greater than the correlation between  $A$  and  $B'$ . He offers a “completely schematic example” to show that some values for the relevant probabilities would satisfy this claim ([Williamson, 2000], 85). Using our notation, and letting  $\Pr(A|B)$  stand for the probability of  $A$  conditional on  $B$ , he must show at least that

$$\Pr(A|K) \gg \Pr(A|B') \tag{2}$$

---

<sup>9</sup>In what follows we modify Williamson’s notation to achieve greater perspicuity. As far as we can see, nothing of philosophical substance is affected by these modifications.

<sup>10</sup>Williamson cautions, “Correlation is itself only one of many explanatory virtues, but it is the one of present interest” ([Williamson, 2000], 83). Yet, correlation *tout court* has no straight-forward explanatory force (cf. [Cartwright, 1983], 21–43). Nevertheless, we are prepared to grant Williamson the connection between correlation and explanation that he needs; we believe there are more serious problems with his argument.

obtains in some plausible cases ([Williamson, 2000], 85).<sup>11</sup>

It is easy to see why (2) is connected to the case Williamson is making. If it obtains, then  $K$  not only makes  $A$  more likely but also makes it significantly more likely than  $B'$  makes it. This would be some reason to think that  $K$  will be a better predictor of  $A$  and, given Williamson's assumptions about explanation, that  $K$  better explains  $A$ . Now the question is whether (2) obtains in any interesting cases. Williamson attempts to meet that challenge by providing a trio of examples.

## 2.2 Examples and Evaluation

Williamson asks us to imagine being at home when a man knocks on the door. We consider not replying, hoping that he will go away. Will he conclude (rightly) that we are hiding from him and be cross with us, or will he suppose that we are not home after all and hold no grudge against us? Williamson suggests, "Whether he would take offence is better predicted by whether he knows than by whether he believes. His taking offence is more highly correlated with knowing that you are in than with believing (truly) that you are in" ([Williamson, 2000], 86). As noted above, a necessary condition for this correlation is that the man's being offended ( $A$ ) is more likely if he knows we are home ( $K$ ) than if he merely has a true belief that we are home ( $B'$ ). Williamson insists that this obtains, suggesting that the man is less likely to revise his beliefs when we fail to reply if he knows than if he merely believes: "someone who knows that you are in has grounds that will not be undermined by your failure to reply" ([Williamson, 2000], 86).<sup>12</sup>

Suppose that the man infers  $p$  from another belief  $q$ . He knocks, but we do not answer. He wonders to himself whether we are at home after all. He ponders how he came to believe  $p$  in the first place. If the man knows  $p$ , then his inference will stand the scrutiny. If he does *not* know  $p$ , then it may be that  $q$  is false or that the inference from  $p$  to  $q$  was fallacious. On reflection, the man may discover that his inference was unsound and give up his belief in  $p$ . The idea is that *knowing*  $p$  precludes a certain way that the man might be moved to reduce his confidence in  $p$ .

We might attempt to answer this by enriching our logical description of

---

<sup>11</sup>Williamson writes that "if the probability of performing the action conditional on knowing  $p$  exceeds the probability of performing it conditional on believing  $p$  truly without knowing  $p$  by much more than the latter exceeds the probability of performing it conditional on failing to believe  $p$  truly, then performing the action is more highly correlated with knowing  $p$  than with believing  $p$  truly" ([Williamson, 2000], 85). That is, he aims to show that  $\Pr(A|K) - \Pr(A|B') \gg \Pr(A|B') - \Pr(A|\neg(B\&p))$ . On the plausible assumption that the right-hand side of that inequality is positive, however, this reduces to (2). Note that correlation is symmetric and as such (2) might be expressed in terms of  $\Pr(K|A)$  (and  $\Pr(B'|A)$ ); however, since Williamson only appeals to correlation in order to show that  $K$  is a causal factor that increases the probability of  $A$ , it makes sense to consider the probability of  $A$  conditional on possible causal factors.

<sup>12</sup>Williamson insists that the difference in revisability is not to be understood as reflecting a difference in confidence, since he does not believe that knowledge is merely a more confident version of belief.

the situation. Let's begin by separating the man's belief when he arrives from his belief after his knock goes unanswered. Let  $B'_0 (K_0)$  be the man's true belief (knowledge) as he first arrives, and let  $B'_1 (K_1)$  be his belief (knowledge) sometime later. If reflection fails to dislodge his belief, then he will be just as cross with us if he truly believes we were home as he would if he knew it; thus

$$\Pr(A|B'_1) = \Pr(A|K_1). \quad (3)$$

However, the assumption of the case was only that he believed or knew we were home as he began to knock. So let  $N$  be the condition that we do not answer, and let  $R$  be the condition that he has reflected on his belief in  $p$ . Supposing him to be similarly confident in each case, there is some degree to which he will be discouraged if we do not answer — this will depend on how likely he thinks we are not to answer if we are home, but the effect should be the same whether he knows or merely believes. From that consideration alone, we would think

$$\Pr(B'_1|N\&B'_0) = \Pr(K_1|N\&K_0). \quad (4)$$

If he believes, though, there is a chance that he believes on the basis of some unsound inference that a moment's reflection will reveal — this is not so if he knows  $p$ . As such,

$$\Pr(B'_1|N\&R\&B'_0) < \Pr(K_1|N\&R\&K_0). \quad (5)$$

It follows on plausible further assumptions<sup>13</sup> that

$$\Pr(A|N\&K_0) > \Pr(A|N\&B'_0). \quad (6)$$

Recall, though, that the left-hand probability in (2) must be *much greater* than the right-hand probability; one may yet worry that the difference between the two sides of (6) is too small — viz., too small to justify moving from explanation in terms of belief exclusively to explanation in terms of knowledge. Unfortunately, we don't see a way of deciding whether this is so without specifying probabilities in an implausibly precise way.

Therefore, let's take another tack and consider the case where the man believes for reasons which will not collapse under scrutiny; let  $B_0^*$  be such a case when he arrives, and let  $B_1^*$  be that case later. The wedge Williamson drives between knowledge and true belief cannot be driven between knowledge and true-belief-that-will-survive-scrutiny. At least in this case,  $K_0$  and  $B_0^*$  will do equivalent explanatory work.<sup>14</sup> Since we can specify  $B^*$  without any refer-

<sup>13</sup>Suppose, for instance, that he is sure to reflect if we do not answer (i.e., that  $\Pr(R|N) = 1$ ) and that  $A$  is only directly influenced by  $B'_1$  and  $K_1$  (i.e., that  $A$  is independent of  $B'_0 (K_0)$  conditional on  $B'_1 (K_1)$ ).

<sup>14</sup>The force of this response to Williamson depends on supposing that knowledge is distinct from true-belief-that-will-survive-scrutiny. Since this supposition follows from Williamson's insistence that knowledge is not analyzable in terms of belief plus other ingredients (chapter 1), it is not open to him to answer our response by denying the supposition. (That said, we believe the supposition in question can be motivated by a number of considerations independent of Williamson's commitments.)

ence to *knowledge*, this first example doesn't establish Williamson's claim that knowledge is indispensable for psychological explanation after all.<sup>15</sup>

It is even possible that  $B^*$  has an explanatory *advantage* over  $K$ . We have supposed so far that in this example  $B'_0$  and  $K_0$  will yield  $A$  in all of the same circumstances, except when  $B'_0$  will not bear up under scrutiny. The asymmetry between  $K_0$  and  $B'_0$  was that  $B'_0$  could fail to produce  $B'_1$  for some range of cases in which  $K_0$  would still produce  $K_1$ . Yet, as Williamson later argues, one can know something and then later cease to know it ([Williamson, 2000], 206, 218–219). Thus, it is possible for our visitor to know we are home when he arrives, but merely to forget or conclude falsely on reflection that we are not. There are some cases in which  $K_0$  would fail to produce  $K_1$ ; there is no reason to suppose anything about whether  $B'_0$  would fail to produce  $B'_1$  in those cases. Williamson draws out the asymmetry in the example that gives  $K$  an explanatory edge over  $B'$ , but other asymmetries might give  $B'$  (and *a fortiori*  $B^*$ ) an explanatory edge over  $K$ . We do not think this objection is fatal, however, since Williamson may easily shift from  $K$  to  $K^*$  and insist that the greatest explanatory work can be done by knowledge-that-will-survive-scrutiny. Regardless, we see no reason to favor  $K^*$  over  $B^*$ .<sup>16</sup>

One may object to our proposal on the grounds that  $B^*$  and  $K^*$  are *ad hoc* and ill-suited to explain action, regardless of how well correlated they are with actions of particular kinds. For, one might insist, an *explanans* will be *ad hoc* unless it is fully specified in terms of what happens earlier than the time of the *explanandum*. But  $B^*$  *qua* “belief-that-will-survive-scrutiny” is given, verbally, in terms of what happens later; the belief will survive scrutiny or it will not. If this is right, then  $K$  is to be preferred over  $B^*$  after all. But we are not moved by this objection. First,  $B^*$ 's survival of scrutiny, although after the incident at your front door, is still prior to the man becoming angry ( $A$ ). We do not see why it is *ad hoc* to explain some future event in terms of what is presently the case and what will come to pass between now and then.<sup>17</sup> Second, nothing in principle precludes  $B^*$  from being expressed in terms of factors present at the time of  $B$ . The factors in virtue of which  $B$  will persist are, after all, already present (and presumably would be unpacked in an account of belief persistence — the need for which is independently motivated). Since the will-survive-scrutiny element of  $B^*$  can thus be taken as elliptical for already present factors (to be filled in by an account of belief persistence), the proposed constraint on explanations gives

<sup>15</sup>Williamson provides the further example of a fox searching the woods for a rabbit (86–87). It is structurally equivalent to the man-at-the-door example, and the responses we have given apply, *mutatis mutandis*.

<sup>16</sup>This answers a potential objection against our account in terms of  $B^*$  — viz., that such an account demands additional apparatus in the form of a theory of belief persistence, and is therefore less simple than a non-dynamical theory in terms of  $K$ . Insofar as Williamson is right that knowledge can fail to persist (a point we are inclined to agree with), knowledge will fail to be explanatory unless it, too, is supplemented with a story about persistence.

<sup>17</sup>Indeed, such explanations may be preferable in virtue of being more complete than explanations given only in terms of what is presently the case. Suppose we ask why the man's knowing we are home ( $K$ ) is apt to make him cross with us ( $A$ ). The explanation is more complete if it explains how he will react to our ignoring his knocking. Why, then, forswear its use?

no reason for preferring  $K$  over  $B^*$  after all.

Williamson also considers an example in which “the predictive difference between knowledge and true belief is mediated by the cultural significance of knowledge” ([Williamson, 2000], 87). Imagine a notorious celebrity who commits a homicide in front of several witnesses. In order to avoid conviction, she tracks them down and kills them one by one. She is interested in killing the people who *know* that she committed the first murder and not the people who merely *believe truly* that she did it. Witnesses who do not actually know will not stand up under cross-examination, Williamson suggests, so she may not feel the need to kill them. Although the sort of scrutiny here is different, this is still a case where true-belief-that-will-survive-scrutiny will do the same explanatory work as knowledge. The witness whose testimony will hold up is a danger to the killer whether he knows or merely believes persistently (i.e., with sufficient persistence that he can convince jury members).

### 3 Conclusion

Williamson urges us to supplement a psychology of belief with a psychology of knowledge. We are unpersuaded. First, as discussed in §1, we do not believe that Williamson’s proposed answers to common motivations for the explanatory centrality of belief are successful. Second, as we have argued in §2, the discussion of probabilities only shows that there are conditions that, if satisfied, would make it the case that the correlation between action and a particular explanans is stronger than the correlation between action and one rival explanans. If we make further concessions about the explanatory power of correlations, then what follows is that there are possible conditions in which knowledge would be more useful than true belief in explaining certain actions. However, even confining ourselves to the explanation of these actions in these conditions, belief-that-will-survive-scrutiny does as well as knowledge. For these reasons, we are inclined not to accept Williamson’s offer.<sup>18</sup>

### References

- [Brueckner, 2002] Brueckner, A. (2002). Williamson on the primeness of knowing. *Analysis*, 62(3):197–202.
- [Cartwright, 1983] Cartwright, N. (1983). *How the Laws of Physics Lie*. Oxford University Press, New York.
- [Fodor, 1975] Fodor, J. A. (1975). *The Language of Thought*. Harvard University Press, Cambridge, Massachusetts.

---

<sup>18</sup>This paper grew out of a reading group on *Knowledge and Its Limits* at UCSD in winter 2002; we are grateful to the other members of that group for discussions of these matters, and to Sara Bernal, Nancy Cartwright, Ram Neta, and Timothy Williamson for helpful comments on earlier drafts.

- [Fodor, 1980] Fodor, J. A. (1980). Methodological solipsism considered as a research strategy in cognitive science. *Behavioral and Brain Sciences*, 3:63–109.
- [Fodor, 1987] Fodor, J. A. (1987). *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. MIT Press, Cambridge, Massachusetts.
- [Jackson, 1996] Jackson, F. (1996). Mental causation. *Mind*, 105:377–413.
- [Kim, 1982] Kim, J. (1982). Psychophysical supervenience. *Philosophical Studies*, 41:51–70. Reprinted in [Kim, 1993], 175–193.
- [Kim, 1993] Kim, J. (1993). *Supervenience and Mind: Selected Philosophical Essays*. Cambridge University Press, New York.
- [McGinn, 1989] McGinn, C. (1989). *Mental Content*. Basil Blackwell, New York.
- [Plantinga, 1993] Plantinga, A. (1993). *Warrant: The Current Debate*. Oxford University Press, New York.
- [Schiffer, 1987] Schiffer, S. (1987). *Remnants of Meaning*. MIT Press, Cambridge, Massachusetts.
- [Stich, 1978] Stich, S. (1978). Autonomous psychology and the belief-desire thesis. *The Monist*, 61:573–591.
- [Williamson, 2000] Williamson, T. (2000). *Knowledge and Its Limits*. Oxford University Press, Oxford.