



## Knowledge and its Limits

Timothy Williamson

<https://doi.org/10.1093/019925656X.001.0001>

Published: 2002

Online ISBN: 9780191598678

Print ISBN: 9780199256563

Search in this book

CHAPTER

## 2 Broadness

Timothy Williamson

<https://doi.org/10.1093/019925656X.003.0003> Pages 49–64

Published: October 2002

### Abstract

That knowing is a mental state is inconsistent with internalism, the claim that whether one is in a given mental state depends only on what is going on inside one's head, for the truth of what one knows may involve the external environment. More familiar kinds of externalism in the philosophy of mind have concerned the content of a mental state; the present view extends externalism to the propositional attitudes to those contents. Attempts to extract belief or rational belief as the internalist, purely mental core of knowledge are refuted. The causal efficacy of knowing is upheld by a consideration of cases in which our best causal explanation of an extended action appeals to knowing rather than believing.

**Keywords:** [action](#), [belief](#), [causal](#), [explanation](#), [externalism](#), [internalism](#), [mental state](#), [propositional attitude](#)

**Subject:** [Epistemology](#), [Metaphysics](#)

**Collection:** [Oxford Scholarship Online](#)

### 2.1 Internalism and Externalism

The thesis that knowing is a mental state faces a further series of challenges. They come from a picture of the mind known in current jargon as *internalism*, in a sense of the term more prevalent in the philosophy of mind than in epistemology.

When I attribute a mental state to you in ordinary language, the implications of my statement can easily outrun your boundaries. I say that you see paper; as every sceptic knows, you could be in the same internal state as someone who sees paper without seeing paper yourself: my statement is true only if paper is there before your eyes, outside you. In some sense, my statement is not purely about you. For theoretical purposes, would it not be more perspicuous to resolve the mixture into its underlying elements, by separating a statement purely about you from another purely about the environment external to you? After

all, causation is local—no action at a distance—so does not the causal explanation of your actions require the isolation of what is local to you from background conditions on the environment? This resolution might amount to an analysis, giving a necessary and sufficient condition for the truth of my original statement. Alternatively, it might replace that statement without being equivalent to it, by doing its causal-explanatory work better. Either way, internal and external factors are separated. Such a picture is internalist.

In the present sense, internalists hold that one's mental states are determined by one's internal physical states; the mind is in the head. As a characterization of internalism, that is not fully general, for it neglects radically dualist versions, but the form of the argument below would be little changed if 'physical state' were replaced by 'phenomenal state', understood as designating states constitutively independent of the environment. For simplicity, we can focus on the currently most popular version of internalism.

p. 50 Internalism provides a deeper motive for denials that knowing is a mental state. For since knowing is factive, whether one knows  $p$  constitutively depends on the state of one's external environment whenever  
↳ the proposition  $p$  is about that environment. Consequently, whether one knows  $p$  is not determined by one's internal physical state. For example, whether one knows that it is raining is not determined by one's internal physical state, for it also depends on the weather. If it is not raining then one does not know that it is raining, whatever one's internal physical state. Thus, for the internalist, knowing is not a mental state. Jerry Fodor drew just such a conclusion from his formality condition, according to which mental states and processes defined over representations apply to them in virtue of the syntax of the representations: 'Since, on that assumption [that you can't know what's not the case], 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*)' (1981: 228). By contrast, an externalist conception frees us to affirm that knowing is a mental state.

The issue ramifies. On the internalist picture, knowing is a metaphysical hybrid, a mixture of mental states with mind-independent conditions on the external world. Even Tyler Burge, who has done as much as anyone to develop an externalist understanding of the mental, writes that factive verbs like 'know', 'regret', 'realize', 'remember', 'foresee' and 'perceive' 'suggest an easy and clearcut distinction between the contribution of the individual subject and the objective, "veridical" contribution of the environment to making the verbs applicable' (1979: 85). Burge wisely adds in parentheses, 'Actually the matter becomes more complicated on reflection'. The internalist naturally tries to break the supposed mixture down into its elements, to analyse knowing in terms of believing, truth, and further factors. Even so-called externalist analyses of knowledge, where the further factors are causal or counterfactual, concede the internalist assumption that believing is somehow more basic than knowing. Thus internalism also provides a deeper motive for attempts to analyse knowing in terms of believing, truth, and other factors. We may call such attempts *the reductionist programme for knowledge*.

Internalists who regard knowing itself as complex do not thereby commit themselves to the same view of the concept *knows*. A simple concept might be defined by ostension of complex exemplars. Thus internalism motivates the reductionist programme for knowledge more strongly at the metaphysical than the conceptual level. The emphasis in this chapter will therefore be on metaphysical rather than conceptual issues.

p. 51 We may assume that all attempts so far to carry out the reductionist programme for knowledge have failed. That suggests that it is misconceived. ↳ Its failure also suggests that internalism itself is misconceived, insofar as it motivates the reductionist programme. A conception of knowing that is thoroughly externalist in the present sense will dispense with the programme. On such a conception, as developed in the previous chapter, knowing is not a metaphysical hybrid, because it cannot be broken down into such elements.

Section 2.2 briefly illustrates the nature of the case for externalism, without attempting to state it in detail. The aim is rather to draw a comparison between more familiar disputes between internalists and externalists, over the contents of propositional attitudes, and the present dispute, over the attitudes to those contents, and to suggest that the case for externalism about mental attitudes is as good as the case for externalism about mental contents. The main target of criticism in sections 2.3 and 2.4 is the idea that there are good grounds for combining externalism about the contents of the attitudes with internalism about the attitudes themselves. One overall argumentative strategy is to show that objections to the involvement of factive attitudes in genuine mental states are sound only if corresponding objections to the involvement of broad contents in genuine mental states are also sound. For example, cases in which a difference in the external environment constitutes a difference in knowing hardly show that knowing is not a mental state, unless cases in which a difference in the external environment constitutes a difference in broad content show that believing a broad content is not a mental state. Externalism about factive mental attitudes is as well placed as externalism about mental content. Chapter 3 will state a deeper case for externalism on both fronts.

## 2.2 Broad and Narrow Conditions

We can define the issues in more rigorous terms to address them more effectively. What exactly is the distinction between the internal and the external? The boundaries of the agent which our attributions of mental states outrun are spatio-temporal boundaries. The spatial boundary is naturally identified with that of the agent's body, although for present purposes it could just as well be identified with that of the brain (or the head). But only what goes on within the agent's body at the time of action counts as internal, for past bodily goings on are not local in the sense in which causation is supposed to be local. The internal will be identified with the total internal physical state of the agent at the relevant time, the external with the total physical state of the external environment.

p. 52 Everything said here will be consistent with the mildly physicalist assumption that the internal and the external are jointly exhaustive as well as mutually exclusive, in the sense that the total internal physical state of the subject and the total physical state of the external environment jointly determine the total state of the world: no difference in the latter without a difference in at least one of the former.

Some terminology will help. A *case* is a possible total state of a system, the system consisting of an agent at a time paired with an external environment, which may of course contain other subjects. A case is like a possible world, but with a distinguished subject and time: a 'centred world' in the terminology of David Lewis (1979). Different cases can distinguish different subjects and times. Whatever is nomically possible counts as 'possible' in the relevant sense; whether anything else does can be left open for present purposes.

A *condition* obtains or fails to obtain in each case. Conditions are specified by 'that' clauses. The pronoun 'one' and the present tense in such clauses refer to the distinguished agent and time respectively. Thus the condition that one is happy obtains in a case  $\alpha$  if and only if in  $\alpha$  the agent of  $\alpha$  is happy at the time of  $\alpha$ .

A condition C *entails* a condition D if and only if for every case  $\alpha$ , if C obtains in  $\alpha$  then D obtains in  $\alpha$ . The conditions C and D are identical if and only if for every case  $\alpha$ , C obtains in  $\alpha$  if and only if D obtains in  $\alpha$ . Truth-functions of conditions are defined in the obvious way; for example, the conjunction of C and D obtains in  $\alpha$  if and only if both C and D obtain in  $\alpha$ . The criterion of identity for conditions ensures that such truth-functions have unique values.

A case  $\alpha$  is *internally like* a case  $\beta$  if and only if the total internal physical state of the agent in  $\alpha$  is exactly the same as the total internal physical state of the agent in  $\beta$ . A condition C is *narrow* if and only if for all cases  $\alpha$  and  $\beta$ , if  $\alpha$  is internally like  $\beta$  then C obtains in  $\alpha$  if and only if C obtains in  $\beta$ . In other terminology, narrow

conditions supervene on or are determined by internal physical state: no difference in whether they obtain without a difference in that state. C is *broad* if and only if it is not narrow. A state S is narrow if and only if the condition that one is in S is narrow; otherwise S is broad. Internalism is the claim that all purely mental states are narrow; externalism is the denial of internalism.

p. 53 When we attribute mental states to each other in ordinary language, the conditions of which we speak are often broad. That one sees Naples, that one remembers Naples, that one keeps referring to Naples—all are broad conditions, because none obtains in cases in which one lacks even indirect causal connection with Naples, whereas one's internal physical state has no such necessary dependence on a city. Similarly, that one loves Mary and that one hates Mary are broad conditions, for they depend on a relation to the particular individual named.

The semantics of ascriptions of content to propositional attitudes in natural languages is a notorious source of broad conditions. In retrospect we can trace the idea back to Hilary Putnam (1973), as interpreted in the light of Burge's work, such as his 1979, 1986a, and 1986b. For example, the sentence 'One believes that there are tigers' expresses a broad condition. To check that it does, consider a counterfactual world like ours except that the only tiger-like creatures, although similar in appearance to tigers, are quite different in evolutionary ancestry and internal constitution. The differences are in respects about which ordinary non-zoologists are ignorant. Call the tiger-like creatures *schmigers*. Clearly, schmigers do not belong to the same species as tigers; they are not tigers. In the counterfactual world I have a *doppelgänger*, twin-TW, who is in exactly the same internal physical state as I am. I believe truly that there are tigers. I express my belief by saying 'There are tigers'. Twin-TW expresses his belief by saying 'There are tigers', too. If he believes that there are tigers then he is wrong, for in his circumstances there are no tigers; there are only schmigers. But twin-TW is no more mistaken on this matter than I am; both of us are ignorant rather than mistaken about those specific features that differentiate tigers from schmigers. Since twin-TW believes truly, he does not believe that there are tigers. Rather, his belief is true if and only if there are schmigers. Thus I differ from twin-TW in believing that there are tigers, even though we are in exactly the same internal physical state. John McDowell (1977) and Gareth Evans (1982) identify a similar phenomenon in relation to singular thoughts. I believe that this screen flickers; someone could be in exactly the same total internal physical state without believing that this screen flickers, because what (if anything) he believes flickers is not this screen but another with the same appearance in front of him. Thus the condition that one believes that this screen flickers is broad. Similar arguments apply to a vast range of contents, and of attitudes to those contents. We may say derivatively that a content is broad if, for every attitude, the condition that one takes that attitude to that content is broad. Most contents ascribed in natural language are broad.

p. 54 The internalist is obliged to concede that content ascriptions in natural languages express broad rather than narrow conditions, but nevertheless insists that they consequently fail to reflect the structure of the underlying facts. On this view, such ascriptions characterize the subject by reference to a mixture of genuinely mental states and conditions on the external environment. The challenge to such an internalist is to make good this claim by isolating a level of description of contentful attitudes that is both narrow and genuinely mental, not merely neuro-physiological. If there is such a mixture of the internal and the external, it should be possible to separate out its constituents. The broadness of content ascriptions in natural languages shows that the required level of description does not simply lie to hand, but must be constructed; its effect is therefore to put the burden of proof on the internalist.

Parallel considerations apply to internalism about the attitudes themselves. Factive propositional attitudes are a source of blatantly broad conditions, whether or not their contents are broad. Even when the sentence 'One believes that A' does not express a broad condition, the conditions expressed by 'One knows that A', 'One sees that A', and 'One remembers that A' are almost always broad. While conceding this, the internalist nevertheless insists that such constructions fail to reflect the structure of the underlying facts. Factive constructions are held to characterize the subject by reference to a mixture of genuinely mental states and

conditions on the external environment. As before, the challenge to the internalist is to make good this claim by isolating a level of description that is both narrow and genuinely mental. The effect of the broad natural language semantics is again to put the burden of proof on the internalist. On the view developed in Chapter 1, the factive states are as genuinely mental as any states are. The internalist disagrees, and tries to find a narrow non-factive attitude that exhausts the mental reality underlying the broad factive attitude. The next section examines such attempts; they all prove to be inadequate.

## 2.3 Mental Differences Between Knowing and Believing

The argument from internalism to the denial that knowing is a mental state can now be stated in more detail. First, assume that knowing is a mental state:

- (1) For every proposition  $p$ , there is a mental state  $S$  such that in every case  $\alpha$ , one is in  $S$  if and only if one knows  $p$ .

Given (1), any difference in knowing involves a difference in mental states. That is, (1) entails that knowing  $p$  supervenes on one's (total) mental state in this sense:  $\hookleftarrow$

- (2) For all propositions  $p$  and cases  $\alpha$  and  $\beta$ , if one is in exactly the same mental state in  $\alpha$  as in  $\beta$ , then in  $\alpha$  one knows  $p$  if and only if in  $\beta$  one knows  $p$ .

The argument from (1) to (2) is immediate if one defines 'one is in exactly the same (total) mental state in  $\alpha$  as in  $\beta$ ' as 'for all mental states  $S$ , in  $\alpha$  one is in  $S$  if and only if in  $\beta$  one is in  $S$ '. Whether, conversely, (2) entails (1) depends on whether what supervenes on one's mental state is itself a mental state, a question which need not be settled here. Statement (2) could also stand on its own, without such an analysis of the equivalence relation of exact sameness of mental state; then, unlike (1), it would involve no commitment to an ontology of mental states or any consequent problems in individuating states.

Whether or not it is derived from (1), (2) is commensurable with the internalist premise that one's mental state supervenes on one's physical state, in other words, that the condition that one is in a given mental state is narrow:

- (3) For all cases  $\alpha$  and  $\beta$ , if  $\alpha$  is internally like  $\beta$ , then one is in exactly the same mental state in  $\alpha$  as in  $\beta$ .

Together, (2) and (3) entail that knowing  $p$  supervenes on one's internal state, for supervenience is transitive:

- (4) For all propositions  $p$  and cases  $\alpha$  and  $\beta$ , if  $\alpha$  is internally like  $\beta$ , then in  $\alpha$  one knows  $p$  if and only if in  $\beta$  one knows  $p$ .

According to (4), the condition that one knows  $p$  is narrow. But (4) is uncontroversially false. Of two people in exactly the same internal physical state, one may know that it is raining while the other, as the result of an elaborate hoax, believes falsely that it is. One can know  $p$ , all but for the state of the environment, without knowing  $p$ , in the sense that one can be in exactly the same internal physical (not: mental) state as someone who knows  $p$  without oneself knowing  $p$ . Since internalists accept (3), they deny (2), the other premise from which (4) was deduced. Since (1) entails (2), they also reject (1).

Having denied that knowing is a mental state, internalists naturally seek to factorize it into mental and non-mental components. Since believing may appear to be determined by one's internal physical states, and therefore to qualify as a mental state by internalist lights, it is an obvious candidate constituent. The idea that the mental (or psychological) component of knowing is simply believing seems to be expressed in a remark by Stephen Stich, endorsed by Jaegwon Kim: 'what knowledge adds to belief is psychologically

p. 56 irrelevant' (Stich 1978: 574, quoted in  $\hookleftarrow$  Kim 1993: 188; Kim 1993: 175–93 is a clear statement of the kind of view opposed here). Because believing is such an obvious candidate, even those who concede externalism about mental content may be inclined to internalism about the attitude of knowing, regarding it as a mixture of mental and non-mental elements.

In present terms, the claim that knowing  $p$  adds nothing mental to believing  $p$  comes to this:

- (5) For all propositions  $p$  and cases  $\alpha$ , if in  $\alpha$  one believes  $p$  then in some case  $\beta$  one is in exactly the same mental state as in  $\alpha$  and one knows  $p$ .

For if (5) is false, one can believe  $p$  while in a total mental state  $T$  incompatible with knowing  $p$ ; but then the information that one knows  $p$  adds something mental to the information that one believes  $p$ , for it implies that one is in a total mental state other than  $T$ . Thus if knowing  $p$  adds nothing mental to believing  $p$ , then (5) holds. Conversely, if (5) holds, then knowing  $p$  imposes no constraints on one's mental state beyond those already imposed by believing  $p$ , so knowing  $p$  adds nothing mental to believing  $p$ .

Now (5) implies that knowing  $p$  is not a mental state, given that not knowing  $p$  is compatible with believing  $p$ . For if knowing  $p$  is a mental state, then anyone in exactly the same mental state as someone who knows  $p$  also knows  $p$ . More precisely, (1) formalizes the claim that knowing  $p$  is a mental state; (1) entails (2); (2) and (5) entail that in every case if one believes  $p$  then one knows  $p$ . Since not knowing  $p$  is compatible with believing  $p$ , knowing is a mental state only if (5) is false.

However, (5) fails, for reasons independent of internalism. One kind of case involves false propositions about the subject's own mental state. For example, let  $p$  be the proposition that someone is alert. Suppose that in case  $\alpha$  one falsely believes that someone is alert solely on the basis of one's false first-person present-tense belief that one is alert. If in case  $\beta$  one is in exactly the same mental state as in  $\alpha$ , then in  $\beta$  one also believes that someone is alert solely on the basis of one's first-person present tense belief that one is alert; since one's level of alertness is itself a feature of one's mental state, then in  $\beta$  one is not alert, so one's belief that one is alert is false; since that false belief is the only basis for one's belief that someone is alert, one does not know that someone is alert. Thus (5) fails. Counter-examples also occur when someone believes a necessarily false proposition. It is possible falsely to believe that  $79 + 89 = 158$ ; it is impossible to know that  $79 + 89 = 158$ . These examples do not depend on any externalist assumptions about the contents of the beliefs.

p. 57 Since those counterexamples involve false beliefs, the internalist  $\hookleftarrow$  might suppose the remedy to lie in the revised claim that knowing  $p$  adds nothing mental to believing  $p$  truly:

- (6) For all propositions  $p$  and cases  $\alpha$ , if in  $\alpha$  one believes  $p$  truly then in some case  $\beta$  one is in exactly the same mental state as in  $\alpha$  and one knows  $p$ .

Statement (6) implies that knowing  $p$  is not a mental state, given that not knowing  $p$  is compatible with believing  $p$  truly.

Even (6), however, is subject to counterexamples which are independent of externalism. Someone may believe truly that garlic is healthy to eat for reasons so confused and irrational as to be incompatible with knowing that garlic is healthy to eat; since his confusion and irrationality is an aspect of his mental state, no one could be in exactly the same mental state and know that garlic is healthy to eat. Attempts to rule out such cases will be plagued by notorious difficulties in stating a correct justification condition on knowledge (see Shope 1983: 45–118).

Even if (6) had been defensible, it would not have solved the original problem, which was, in internalist terms, to isolate the mental component of knowing, to say what mental state knowing adds nothing mental

to. By specifying that the belief be true, (6) fails to do that. This latter objection is not met by a justification condition added to (6).

Believing  $p$  is in any case too unspecific a state to constitute the mental component of knowing  $p$ . Knowing  $p$  excludes: believing  $p$  solely for sufficiently confused and irrational reasons. The supposed narrow mental component of knowing  $p$  must include not just believing  $p$  but doing so without those kinds of confusion and irrationality. We must therefore consider the suggestion that the mental component of knowing is *rationally* believing (Fricker 1999). 'Rationally' here need not imply the ability to articulate reasons, but only the avoidance of irrationality; languageless animals and young children may still count as knowing. In place of 'rationally believes' we could also write 'has a justified belief'. Now if rationally believing is the mental component of knowing, then the latter adds nothing mental to the former:

(7) For all propositions  $p$  and cases  $\alpha$ , if in  $\alpha$  one rationally believes  $p$  then in some case  $\beta$  one is in exactly the same mental state as in  $\alpha$  and one knows  $p$ .

p. 58 Now (7) may also escape the original counterexamples to (5), for the internalist might count as not believing rationally the subject who believes that  $79 + 89 = 158$  or that someone is alert solely on the basis of a false belief that he is alert. Moreover, (7) implies that knowing  $p$  is not  $\hookrightarrow$  a mental state, given that not knowing  $p$  is compatible with rationally believing  $p$ .

Perhaps (7) looks congenial to externalism about the contents of one's attitudes but not to externalism about one's attitudes to those contents. It is not. Suppose that it looks and sounds to me as though I see and hear a barking dog; I believe that a dog is barking on the basis of the argument 'That dog is barking; therefore, a dog is barking'. Unfortunately, I am the victim of an illusion, my demonstrative fails to refer, my premise sentence thereby fails to express a proposition, and my lack of a corresponding singular belief is a feature of my mental state, according to the content externalist. If I rationally believe that a dog is barking, then by (7) someone could be in exactly the same mental state as I actually am and know that a dog is barking. But that person, too, would lack a singular belief to serve as the premise of the inference, and would therefore not know that a dog is barking. Contrapositively, according to (7), I do not rationally believe that a dog is barking, even though there need be nothing internal wrong with my thought processes. Consequently, if the contents of beliefs depend like that on the external environment, then so too does the attitude of rational belief to a given content. In brief, (7) combined with content externalism makes rational belief an externalist mental attitude. If taking the externalist attitude of rational belief to a given content can contribute to one's mental state, why cannot taking the externalist attitude of knowledge to that content also contribute to one's mental state? The combination of (7) and content externalism makes the denial that knowing is a mental state ill-motivated.

My belief that a dog is barking may easily be true in the example, so to replace 'rationally' by 'rationally and truly' would gain nothing.

p. 59 Indeed, (7) faces a further problem, one independent of content externalism. We could make (7) trivially true by *defining* 'in case  $\alpha$  one rationally believes  $p$ ' as 'in some case  $\beta$  one is in exactly the same mental state as in  $\alpha$  and one knows  $p$ '. If knowing  $p$  entails believing  $p$  and believing  $p$  is a mental state, then such a definition would ensure that believing  $p$  was a necessary condition for rationally believing  $p$ . But it would neither isolate the mental component of knowing in independent terms nor provide any reason to suppose the mental component to fall short of knowing itself. If (7) is to give positive support to the hybrid conception of knowing as a mixture of mental and non-mental components, as (5) was supposed to do, then we should be able to grasp the relevant concept of rationality independently of grasping the concept of knowledge. Can we? Consider a case  $\alpha$  in which one believes that ticket #666 will not win the lottery solely on the basis that its probability of  $\hookrightarrow$  winning is only one in a million. In any case  $\beta$  in which one is in the same mental state as in  $\alpha$  one believes that ticket #666 will not win the lottery only on the same

probabilistic grounds; thus in  $\beta$  one does not know that ticket #666 will not win the lottery. If one had known that the ticket would not win, one would not have bought it. Consequently, by (7), in  $\alpha$  one does not rationally believe that the ticket will not win the lottery. But in  $\alpha$  one's belief is not irrational in any obvious sense independent of considerations of knowledge. It is based on relevant reasons; the problem is just that they are not of a kind that would permit the belief to constitute knowledge. Chapter 8 will argue that considerations of rational belief depend on considerations of knowledge.

Given the failure of (5)–(7), someone might try to capture the idea that the difference between knowing and believing is not mental in the claim that *not knowing*  $p$  adds nothing mental to believing  $p$ . By analogy with (5), the claim is formalized thus:

- (8) For all propositions  $p$  and cases  $\alpha$ , if in  $\alpha$  one believes  $p$  then in some case  $\beta$  one is in exactly the same mental state as in  $\alpha$  and one does not know  $p$ .

For if (8) is false, someone can believe  $p$  while in a total mental state  $T$  incompatible with not knowing  $p$ ; but then not knowing  $p$  adds something mental to believing  $p$ , as the former but not the latter is sufficient, given that one believes  $p$ , for being in a total mental state other than  $T$ . Thus if not knowing  $p$  adds nothing mental to believing  $p$ , then (8) holds. Conversely, if (8) holds, then not knowing  $p$  imposes no constraints on one's mental state beyond those already imposed by believing  $p$ , so not knowing  $p$  adds nothing mental to believing  $p$ .

Now (8) implies that knowing  $p$  is not a mental state, given that knowing  $p$  is compatible with believing  $p$ . For if knowing  $p$  is a mental state, then anyone in exactly the same mental state as someone who knows  $p$  also knows  $p$ . More precisely, (1) formalizes the claim that knowing  $p$  is a mental state; (1) entails (2); (2) and (8) entail that in every case if one believes  $p$  then one does not know  $p$ . Since knowing  $p$  is compatible with and perhaps entails believing  $p$ , knowing is a mental state only if (8) is false.

p. 60 However, (8) is implausible even from an internalist perspective. For example, the proposition  $p$  may concern the subject's own mental state, or be a necessary truth. Internalists will classify direct awareness that one is in pain as a mental state, holding it to depend on nothing external. Presumably, being directly aware that one is in pain is sufficient for both knowing and believing that one is in pain. Thus if one is directly aware that one is in pain, one believes that one is in pain, and  $\downarrow$  could not be in exactly the same mental state without being directly aware, and therefore knowing, that one is in pain. Consequently, (8) fails. Similarly, internalists will classify grasping a proof that  $79 + 89 = 168$  as a mental state, holding it to depend on nothing external. Presumably, they will also hold it to be sufficient for both knowing and believing that  $79 + 89 = 168$ . Thus if one grasps a proof that  $79 + 89 = 168$ , then one believes that  $79 + 89 = 168$ , and could not be in exactly the same mental state without grasping the proof that  $79 + 89 = 168$ , and therefore knowing that  $79 + 89 = 168$ . Again, (8) fails. From a strongly externalist perspective, direct awareness that one is in pain may depend on something external; for example, it may depend on the use of a word in the language community as a whole to mean *pain* rather than something more specific that excludes one's current sensation. Similarly, grasping a proof that  $79 + 89 = 168$  may depend on the mathematical practice of the community as a whole. But for the full-blooded externalist, these external dependencies do not suggest that being directly aware that one is in pain and grasping a proof that  $79 + 89 = 168$  are not mental states. They can stand as counterexamples to (8). A defender of (8) would have to take an intermediate position, on which knowing  $p$  always has non-trivial external necessary conditions not constitutive of the subject's mental state, no matter how trivial the proposition  $p$ . Although this combination of claims is not obviously incoherent, it also has no obvious motivation; (8) lacks the independent plausibility to provide a good reason not to classify knowing as a mental state.

Since the subject rationally and truly believes  $p$  in the examples that are problematic for (8), qualifying 'believes' by 'rationally' or 'truly' in (8) would not help.



The supposed mental component of knowing short of knowing itself is a postulate of philosophical theory, not something provided by our understanding of the relations between knowing and believing. We have no good reason to accept the theory which makes that postulate. The internalist has no head start in the attempt to fence off the mental implications of knowing. That is not yet to say that the attempt cannot succeed; just that we have no reason independent of any case for internalism in general to expect it to succeed.

## 2.4 The Causal Efficacy of Knowledge

---

p. 61 One motive for internalism is the combination of the idea that genuine states are causally efficacious with the idea that mental states are causally  $\hookrightarrow$  efficacious only if narrow. No action at a distance: causation is viewed as local, involving only narrow mental states. Since the property of judging that there is a tiger ahead is broad (because its content is broad), such an internalist denies that the property is causally efficacious, locating causal efficacy in properties that supervene on the subject's internal physical state. After all, they determine the subject's immediate physical movements. Similarly, the internalist will deny that the broad state of knowing that there is a dangerous animal ahead is causally efficacious, locating causal efficacy in the supposedly narrow state of believing that there is a dangerous animal ahead, the state which the knower shares with the victim of a sceptical scenario. For example, according to Harold Noonan (1993: 291–2), knowledge 'is best regarded not as a psychological state, but as a complex consisting of a psychological state (belief) plus certain external factors—not because its status as knowledge is causally irrelevant in action explanation, but because it does not have to be cited, as such, in the psychological explanation of action at all'.

Much needs to be probed and questioned in these internalist ideas. We should not assume that the notion of causal efficacy is clear, or derived from fundamental science, or known to apply only to local connections. Nevertheless, suspicion is legitimate of a purported mental state, reference to which never plays an essential role in causal explanation. In the case of broad contents, a standard externalist move is to argue that attributions of them do play an essential role in causal explanations whose explananda are themselves characterized in broad terms. For example, a hunter shoots a tiger while his counterfactual *doppelgänger* shoots a schmiger. These descriptions of the actions are not capricious, for they are the very ones under which the actions were intended, and therefore the ones to be used when we are trying to see how the actions made sense from the subjects' point of view. Since our explanandum is that the hunter shot the tiger, our explanans will naturally involve broad descriptions of him not true of his *doppelgänger*, such as 'He believed that shooting a tiger would make him popular'.

p. 62 Similar considerations apply to the role of factive attitudes in causal explanation. Consider a causal explanation as simple as 'He dug up the treasure because he knew that it was buried under the tree and he wanted to get rich'. Note that the explanandum ('He dug up the treasure') makes reference to objects in the environment (the treasure) as well as to the subject's immediate physical movements. The internalist cannot substitute 'believe' for 'know' in the explanation without loss, for the revised explanans, unlike the original, does not entail that the treasure was where he believed it to be; the connection between explanans and  $\hookrightarrow$  explanandum is therefore weakened. The explanans does less to raise the probability of the explanandum. As usual, the internalist may react by substituting 'believe truly' for 'believe'. The new explanation is 'He dug up the treasure because he believed truly that it was buried under the tree and he wanted to get rich'. That may be deemed as satisfactory as the original explanation, although it sounds much less natural. However, even the substitution of 'believe truly' for 'know' sometimes involves explanatory loss.

A burglar spends all night ransacking a house, risking discovery by staying so long. We ask what features of the situation when he entered the house led to that result. A reasonable answer is that he knew that there was a diamond in the house. To say just that he believed truly that there was a diamond in the house would be to give a worse explanation, one whose explanans and explanandum are less closely connected. For one possibility consistent with the new explanans is that the burglar entered the house with a true belief that there was a diamond in it derived from false premises. For example, his only reason for believing that there was a diamond in the house might have been that someone told him that there was a diamond under the bed, when in fact the only diamond was in a drawer. He would then very likely have given up his true belief that there was a diamond in the house on discovering the falsity of his belief that there was a diamond under the bed, and abandoned the search. In contrast, if he *knew* that there was a diamond in the house, his knowledge was not essentially based on a false premise. Given suitable background conditions, the probability of his ransacking the house all night, conditional on his having entered it believing truly but not knowing that there was a diamond in it, will be lower than the probability of his ransacking it all night, conditional on his having entered it knowing that there was a diamond in it. It follows that the probability of his ransacking the house all night, conditional on his having entered it believing truly that there was a diamond in it, is lower than the probability of his ransacking it all night, conditional on his having entered it knowing that there was a diamond in it. In this case, the substitution of 'believe truly' for 'know' weakens the explanation, by lowering the probability of the explanandum conditional on the explanans. The substitution of 'believe' without 'truly' for 'know' would do even worse. The argument does not assume that lowering the probability of the explanandum conditional on the explanans strictly entails loss of explanatory power: just that it results in such a loss when, as here, there are no compensating gains.

p. 63

One might be puzzled for a moment by the thought that, in the circumstances, the burglar's true belief constituted his knowledge. Were the effects not the same whatever one calls it? However, this thought does not address the original problem, which concerned the causal efficacy of a *general* state. Different people can share the state of knowing that there was a diamond in the house; this state cannot be equated with, since it is not necessary for, believing truly that there was a diamond in the house. No doubt the particular circumstances that in some sense realize the state in a given case can be described in many different ways; what matters is how relevant those descriptions are to an understanding of the effect in question. It emerged above that the description 'knows *p*' is sometimes more relevant than the description 'believes truly *p*'.

In order to *prove* that reference to states of knowing is essential to the power of a causal explanation, one would need to show that it could not be eliminated in favour of any combination of believing, truth, and so on. There are infinitely many potential substitutes which might be proposed. All that can be done here is to sketch a general strategy for dealing with them; we must not expect to prove that the strategy cannot fail. Given a potential substitute for 'knows', suppose that it does not provide a necessary and sufficient condition for knowing. One then constructs possible cases in which the failure of necessity or sufficiency makes a causal difference, making the proposed substitute not even causally equivalent to knowing. The potential substitute avoids this problem only if it does provide a necessary and sufficient condition for knowing. Thus the search for a substitute for knowing in causally explanatory contexts is forced to recapitulate the history of attempts to analyse knowing in terms of believing, truth, and so on, a history which shows no sign of ending in success.

For example, the substitution of 'believe truly without reliance on false lemmas' for 'know' can bring causal-explanatory loss. Variants of the previous case can be constructed in which the burglar enters the house believing truly that there is a diamond in it without reliance on false lemmas, yet fails to know in virtue of misleading evidence which he does not then possess, but may discover in the course of his search, in which case he will abandon the search. The argument for explanatory loss runs as before. Although knowing is not invulnerable to destruction by later evidence, its nature is to be robust in that respect.

Stubbornness in one's beliefs, an irrational insensitivity to counterevidence, is a different kind of robustness; it cannot replace knowing in all causal-explanatory contexts, for the simple reason that those who know  $p$  often lack a stubborn belief in  $p$ . The burglar's beliefs need not be stubborn. Similarly, he need not feel certain of them; subjective certainty cannot always replace knowing. The same applies to believing truly on the best possible evidence, for the example can be so constructed that the burglar's evidence, although good, is not the best possible. When one works through enough examples of this kind, it becomes increasingly plausible that knowing can figure ineliminably in causal explanations. It is causally efficacious in its own right if any mental state is (see Pettit 1986 and Child 1994: 204–16 for related discussion).<sup>1</sup>

This chapter has stated a preliminary case for externalism about both mental contents and factive mental attitudes to those contents. The next chapter deepens the case and places it in a wider context.

## Notes

---

- 1 Hyman 1999 argues plausibly for another connection between knowledge and action: one knows that  $A$  if and only if one's reason for doing something can be that  $A$ . But it does not follow that one can *explain* knowing that  $A$  as being able to do things for the reason that  $A$  (as Hyman wishes to do). Someone in a Gettier case who believes truly that  $A$  without knowing that  $A$  cannot do  $X_1$  for the reason that  $A$ , and cannot do  $X_2$  for the reason that  $A$ , . . . . But a single failure to know explains all these incapacities. If the incapacities constituted the failure to know, the correlation between the incapacities would be an unexplained coincidence.