

Reflections on Knowledge and its Limits*

Gilbert Harman[†]

Princeton University

June 9, 2003

Williamson's *Knowledge and its Limits* is the most important philosophical discussion of knowledge in many years. It sets the agenda for epistemology for the next decade and beyond.

The main thesis of the book is “knowledge first.” One aspect of this thesis is that knowledge is a basic mental state, a propositional attitude that cannot be explained in terms of belief plus certain other conditions. Given that knowledge counts as a mental state, it follows that people who are internal physical duplicates can be in different mental states. Since a person in a “good” normal situation of veridical perception knows things that are not known by a person who is internally a duplicate but externally

*Timothy Williamson, *Knowledge and its Limits*, (Oxford: Oxford University Press, 2000), pp. xi, 340.

[†]I am indebted to Elizabeth Harman and Mark Greenberg for very useful extensive comments on earlier drafts.

in a “bad” sceptical situation like being a brain-in-a-vat or being perfectly deceived by an evil demon, the person in the good situation has different mental states from the internal duplicate in the bad situation. Williamson rejects what he calls “the internalist conception of mind”.

Williamson argues for several further points. Psychological explanations that appeal to what a person knows cannot always be replaced by equally good explanations in terms of beliefs and environmental conditions. Our mental states are not “luminous,” that is, we do not always know whether or not we are in a given mental state. Knowledge typically involves a “margin for error.” Knowing that one knows typically involves an additional margin and so the *KK*-thesis is false: one might know without knowing one knows. The margin for error condition on knowledge is to be distinguished from a tracking or counter-factual sensitivity requirement.

More speculatively, Williamson identifies a person’s total evidence with what the person knows, $E = K$. Since a person in a bad brain in a vat or evil demon situation does not know as much as a person in the corresponding good situation in which things are as they seem, two people who are internally the same can fail to have the same evidence. Since what it is rational or reasonable to believe depends on one’s evidence, what it is rational to believe in the bad sceptical situation differs from what it is rational to

believe in the good normal situation. Evidential probability is conditional probability on one's evidence, so everything one knows has a probability of one.

Although some of these claims may seem quite counter-intuitive, Williamson defends them and other ideas with the highest level of philosophical argument. I cannot in my remarks do justice to all or even a sizable number of Williamson's themes and arguments. Instead I will take up three issues. First, I want to argue that, even if Williamson is right in his main arguments, there is a sense in which mental states or events are "internal." Second, I want to compare Williamson's idea that one's justified beliefs rest on a foundation of everything one knows with the view defended by Goodman, Rawls, and others that justification starts from where one is, with all of one's current beliefs, intuitions, etc. Third, I want to argue that it fits better with Williamson's overall aim to treat factive attitudes as attitudes toward facts rather than as attitudes toward true propositions.

Is the mental internal?

Williamson argues that knowledge is as good a mental state as belief, desire, and feeling pain. One natural objection is that standard mental states are internal states of a person, whereas whether one knows depends also on the

environment. In response, Williamson rejects “the internalist conception of mind”, noting among other things that the widely accepted “content externalism” about beliefs and desires implies that what beliefs and desires one has can also depend on external things (6). But, without getting into the details of his argument, I would like to point out that two different things might be meant by “the internalist conception of mind.”

Consider mental events and states of a particular person W . On the one hand, these events and states seem to be completely internal to W . On the other hand, “content externalism” implies that mental events and states do not always supervene on states and events internal to the person. There is an apparent conflict here, but is it real?

I suggest that if we distinguish particular occurrences of mental states and events from types or categories of mental states and events, then to speak loosely we can accept internalism for the occurrences while rejecting it for the categories.

Suppose that mental events are particular dated located occurrences and that “internal to W ” means *located entirely within W 's body*. Then it is plausible to me that W 's mental events so conceived are completely internal to W in that sense and are perhaps even realized by physical events in W 's body.

Consider an occurrence of a thought of W 's that it would be nice to have a drink of water. That particular occurrence can be completely internal to W even if its correct categorization as an instance of the *thought that it would be nice to have a drink of water* depends on factors external to W because whether it is a thought about water or about something else depends on whether W 's world is full of water or something else.

Similarly, suppose mental states are also particular dated located items. Williamson himself does not always suppose this. At one point he says, "the idea of a token state is of doubtful coherence" (40), appealing to Steward (1997). At another point, mentioned below, he seems to allow such a supposition, although I may be reading something into the text that he does not intend.

In any event, just for the sake of argument, let us make the assumption that in addition to types of mental state, there are also particular dated located instances of mental states. Given that assumption, it is plausible that W 's mental states so conceived are completely internal to W 's body and perhaps to be identified with or realized by physical states in W 's body. A particular instance of a belief of W 's that Albert Einstein was a physicist is then internal to W even if its correct categorization as an instance of the *belief that Albert Einstein was a physicist* depends on factors external to W

(because whether it is a belief about Einstein or someone else depends on something about the historical source of that instance).

In what follows, I will use the term “occurrence” for any such particular instance of a state or event. Parsons (1990) elegantly defends the appeal to occurrences states and events in this sense as part of an account of the semantics of English.

Next consider an occurrence of W 's knowing that his neighbor is called “Harry.” It is plausible that this occurrence is internal to W , although its correct categorization as an instance of *knowing that one's neighbor is called “Harry”* depends on factors external to W , such as whether W 's neighbor is called “Harry.”

Consider an occurrence of W 's perceiving Harry approaching. At least two different events might be considered occurrences of this perceiving, one more inclusive than the other. There is an event of visual awareness completely internal to W . There is also a larger event containing that internal occurrence plus Harry's approaching plus certain connections between these parts. Either event might be categorized as an occurrence of perceiving Harry approaching. Plausibly, the more narrowly circumscribed completely internal event is an entirely mental occurrence, whereas the larger event is not.

Similar issues arise in the theory of action. Is an act opening the door a purely mental event? What is included in that act? Perhaps there is a purely internal occurrence of an act of will, which might be described in terms of its consequences for bodily movement and the door's opening. If that's the act of opening the door, it is a purely mental occurrence. On the other hand, if the act is identified with a larger event that includes the ensuing bodily movements, perhaps it is not a purely mental event. If the act is identified with an even larger event that includes the door's opening, it is clearly not entirely mental.

So, there is at least one "internalist conception of mind" that as far as I can see is not touched by Williamson's arguments. I mention this, not as an objection to Williamson, but as an aid to understanding his position. Resistance to this aspect of Williamson's view (also McDowell, 1994, 1995) might arise from a failure to distinguish the more plausible internalism of mental occurrences from the rejected supervenience of mental categories on internal states. (I believe that this widely felt resistance confirms the naturalness of supposing that there are particular occurrences of belief and knowledge, despite Steward's claim that there is no natural basis for such an idea. However, I cannot here give this issue the attention it deserves.)

If we assume with Williamson that knowing entails believing, then when-

ever there is an occurrence of knowing that p there is also an occurrence of believing that p . How are these occurrences related? Perhaps they are the same occurrence. That is, perhaps all occurrences of knowing that p are also occurrences of believing that p . This of course does not imply that all occurrences of believing that p are also occurrences of knowing that p . Nor does it imply that the concept *knowing* can be usefully analyzed in terms of the concept *believing* and other concepts.

As I mentioned, Williamson holds that the concept *knowing* cannot be analyzed into more basic concepts. He does endorse the “proposal . . . that knowing is the most general factive stative attitude, that which one has to a proposition if one has any factive stative attitude to it at all” (34). But that is not to define knowing in more basic terms. (I will come back to Williamson’s assumption that knowing is an attitude one has to a proposition.)

If believing is a mental state and all occurrences of knowing are occurrences of believing, there is a sense in which knowing is a mental state—any occurrence of knowing is an occurrence of a mental state. Williamson appears to allow for this despite his earlier remark, noted above, that “the idea of a token mental state is of doubtful coherence.” In a discussion of how a causal explanation might appeal to knowledge rather than belief, he says,

“One might be puzzled for a moment by the thought that, in the circumstances, the burglar’s true belief constituted his knowledge. . . . However, this thought does not address the original problem, which concerned the causal efficacy of a general state” (62). Here he seems at least to acknowledge that this might be a natural thought, although he does not actually endorse it.

It is, of course, a further question whether the concept of *knowing* is a mental state concept in the way that *believing* and *desiring* are. As I have already noted, Williamson answers “yes” to this further question. He argues persuasively that *knowing* is a basic and unanalyzable mental state concept.

Justified belief

Some theories of justified belief suppose that *W*’s beliefs must ultimately be justified from *W*’s *foundations*. These foundations constitute *W*’s ultimate grounds. *W*’s justification must ultimately begin with them (or end, depending on which direction one is going). They may consist partly in certain experiences that provide at least prima facie justification for certain beliefs. In some versions the experiences have to be conscious; other versions (Pollock, 1999) allow for unconscious foundations. The foundations might also contain certain basic beliefs that are prima facie justified with-

out needing a justification that appeals to other beliefs. Nonfoundational beliefs require justification in terms of other justified beliefs, justifications that ultimately terminate in foundational beliefs or experiences.

Coherence theories hold that there are no foundations in this sense; all beliefs require justifications that appeal to other justified beliefs.

General foundations theories (Goodman, 1955; Rawls, 1971) suppose that all of one's present beliefs are foundational. The idea is that one starts where one is, trying to eliminate conflicts among one's beliefs and experiences, trying to make one's overall view more coherent, seeing implications of one's opinions for questions one is interested in answering, seeing what one thinks of these implications, and in general aiming at the no doubt impossible ideal of achieving a perfect "reflective equilibrium" in one's opinions.

Coherence theories and general foundations theories agree that all of one's present beliefs and experiences get treated in the same way. On the other hand, *special foundations* theories distinguish among one's present beliefs and experiences, taking some but not others to be foundational. In special foundationalism, everything one is justified in believing must ultimately be justified from a limited set of foundational beliefs and experiences. Cartesian versions of special foundations severely restrict the foundations to present conscious experiences or to beliefs about a priori axioms and the

evidence of current conscious experience.

Williamson's view is a version of special foundationalism that treats everything that one knows as the relevant foundation. One is justified in believing something if and only if so believing is justified on the basis of everything one knows. In Williamson's view, the size of *W*'s foundations falls between the very limited Cartesian foundations and the very extensive general foundations that would include all of *W*'s beliefs.

It is useful to compare what various approaches to justification say about Williamson's "good" non-sceptical and "bad" sceptical scenarios. In the good scenario *W* has a great deal of knowledge. In one bad scenario *W* is internally the same but is fooled by an evil demon. The coherence theory, the general foundations theory, and Cartesian special foundations theories agree that what *W* is justified in believing in the two cases is exactly the same. Certain reliabilist theories (Goldman, 1979) of justification imply that there is a difference in the two cases if the processes leading to belief in the good case are reliable whereas the processes leading to belief in the bad case are generally unreliable. This consequence of those forms of reliabilism has generally been regarded as so implausible as to raise a conclusive objection to the view (Cohen, 1984; Foley, 1985; Ginet, 1985; Pollock, 1984).

Like reliabilism, Williamson's theory implies that what *W* is justified in

believing differs in good and bad cases because W 's foundations—what W knows—is very different in the two cases even though they are internally the same. The implication that there could be such a contrast between the two cases seems highly counter-intuitive to many reflective people. Williamson's argument for the contrast is not that it seems intuitively right but that it allows one to avoid the skeptical conclusion that W is not justified in his beliefs in the good case. His basic argument is quite simple. In the bad case W is not justified in many of his beliefs, because his resources in the bad case are simply inadequate; so, if W is to be justified in his beliefs in the good case, there must be a difference in what W is justified in believing in the two cases, despite their similarity from one perspective.

The weak spot in this argument (as I have presented it) is its assumption that in the bad case W 's resources are inadequate to allowing W to be justified in many of his beliefs. While the inadequacy assumption may be correct given the limited resources of a Cartesian version of a special foundations theory, it is not correct for a general foundations theory that allows appeal to everything W believes. That basis is sufficient to allow W to be justified in the bad case in believing a great many things, because W has many relevant beliefs.

I conclude that we do not yet have material to decide between Williamson's

special foundations theory and the general foundations theory.

One interesting possible difference between these two views of justification is that a general foundations approach might allow rational recovery even from very extensive errors in cases where Williamson's view would not allow rational recovery because there is no initial knowledge on which to base that recovery.

Factive attitudes and propositional attitudes

As Williamson observes (43), surface grammar seems to indicate that factive attitudes are attitudes toward facts, where facts are distinct from propositions (Vender, 1972). In talking about W being aware that Harry is walking toward him, we say that he is aware of the fact that Harry is walking toward him. That is not just to say that W is aware of the proposition that Harry is walking toward him, even if that proposition is true. (It might be to say that W is aware that the proposition is true, or aware of the truth of the proposition, which is different from saying merely that W is aware of the proposition, even if the proposition is true.) To say that W remembers the fact that Harry promised to meet him later today is rather different from saying that W remembers the proposition that Harry promised to meet him later today. Similarly, to say that W perceives certain facts is not to say that

he perceives certain propositions, again even if they are true propositions.

These facts of surface grammar may seem to pose problems for standard views in the theory of knowledge. For example, what can be quantified over in order to express the claim that, if W knows that p , W believes that p and it is true that p ? What W knows are facts, whereas what is true and what is believed are propositions.

Another issue is this: If what one knows are facts, facts are not true propositions, and evidence is at bottom a relation among propositions, how can what W knows be evidence for the truth of certain conclusions?

Williamson responds to these issues by arguing, first, that a metaphysics that distinguishes facts from true propositions is dubious and, second, that surface grammar does not consistently support that distinction either. But both of these arguments are inadequate.

For the first response, namely, that “it is very doubtful that there are any such things as facts other than true propositions” (43), he refers to Williamson (1999) for further discussion. However, that paper is concerned with something else, namely, with the interpretation of a certain “truthmaking principle” according to which, if a proposition is true, then there is a truthmaker for that proposition, where a truth maker for P is a fact F such that, necessarily, if F exists then P is true. Williamson observes that the

defender of truthmakers must suppose there are possible truth makers that don't exist, so in some sense it is not the truthmaker by itself that makes the proposition true but rather the existence of the truthmaker, where existence is what actualities have and not mere possibilities. It is unclear to me why defenders of facts as truthmakers should be particularly worried by this, especially those who identify facts with those possible states of affairs that obtain. Furthermore, the seemingly sharp distinction in ordinary language between propositions and facts is surely not held hostage to one or another philosophical theory about truthmakers.

Williamson's second response appeals to "the propriety of remarks like 'I always believed that you were a good friend; now I know it' and 'Long before I knew those things about you I believed them' suggest that 'believe' and 'know' do take the same kind of object" (43). But these data are quite inconclusive. The pronoun "it" in Williamson's first example is a pronoun of laziness standing in for the phrase "that you were a good friend". The pronoun is not used referentially. Compare the classic example, "The man who gives his paycheck to his wife is better off than the man who gives it to his mistress," where "it" stands in for the words "his paycheck" and does not refer back to the paycheck mentioned earlier.

Williamson's second example, "Long before I knew those things about

you I believed them,” is quite doubtful. What are the “things” referred to? Williamson holds that they are propositions, but the more explicit remark, “Long before I knew those propositions about you, I believed them,” is odd and cannot be what is meant (Moffett, forthcoming, suggests a way round this worry, however.)

In any event, it is hard to see how one could suppose that the extensive linguistic data distinguishing talk of facts from talk of propositions could be so easily rebutted.

Is it a serious problem that knowing that p is knowing the fact that p rather than knowing the proposition that p ? Williamson thinks that the whole issue is entirely terminological. Suppose “what one knows is a fact, what one believes a proposition, where a fact is not a true proposition . . . then knowing is not a *propositional* attitude, and much of the terminology of this book might need revision, although the substance of the account would remain” (43).

To illustrate the sort of terminological revision that might be needed, Williamson says, “Perhaps we could treat ‘that A’ as elliptical for ‘the fact that A’ . . . where the reference of this is presumably determined by the proposition p expressed by ‘A’; it is therefore some function f of p . thus to know that A is to know the $f(p)$, and hence to stand in a complex relation

expressed by ‘know’, ‘the’, and ‘f’ to the proposition expressed by ‘A’. But then with only a slight change of meaning we could use the word ‘know’ for that complex relation to a proposition. Thus, . . . knowing would still involve a propositional attitude” (43). (Somewhat related proposals are discussed by Parsons, 1993; King, forthcoming; and Moffett, forthcoming.)

To avoid confusion let us use the terms *know'* and *knowledge'* for the relevant propositional attitude, so that to know' that p is to be in a relation to the proposition that p rather than the fact that p . When Williamson says that, if we accept the distinction between facts and true propositions, “much of the terminology of the book might need revision,” is he suggesting that his theory would then not be a theory of knowing but of knowing'? Should he then be interpreted as arguing that knowing' is a basic mental relation?

On the contrary, I suggest that it is more in keeping with Williamson's “knowledge first” idea to treat genuinely factive attitudes as more basic than mere propositional attitudes, where genuinely factive attitudes are attitudes to facts not propositions. That approach fits with wide as opposed to narrow functionalist theories of mental states. Versions of wide functionalism are defended by Sellars (1963), Taylor (1964), Dennett (1969), Harman (1973), Bennett (1976), Stampe (1977), Dretske (1981), and Stalnaker (1984). These approaches tend to take the most basic cases of internal

representations to be cases in which nothing goes wrong and so are cases of knowledge. Allowing for errors and other complications allows for more belief like states.

The more solipsistic narrow approach is found e.g. in Quine (1960), early Putnam (1960, 1964, 1967a, and 1967b), Field (1977), Fodor (1980), Loar (1981) and Lycan (1984). One reason for rejecting the narrow approach in favor of the wide approach are that explanations of reactions in terms of hallucination are normally parasitic on explanations of how people react to veridical perception. It is hard to explain hand-eye coordination without invoking the wide story. It's impossible to explain the functions of sense-organs or motor-organs narrowly. Similarly for functional explanations of brakes on a car or the workings of a radio. Methodological solipsism is incoherent in psychology. There is no way to isolate the relevant psychological components of a creature without considering how those components function in enabling the creature to deal with its environment (Harman 1988).

Of course, even if such an argument for "knowledge first" is accepted, it does not follow immediately that knowledge is a more basic concept than belief in our resulting conceptual scheme or even that knowledge is a paradigm mental state. Although people are able to acquire certain concepts of physics, like energy, momentum, mass, weight, and force, only because

they start with more ordinary concepts, it does not follow that the ordinary concepts are concepts of physics or that they are more fundamental in the conceptual scheme of physics.

Concluding Remarks

I have discussed three aspects of Williamson's book

First I suggested that there are two ways in which mental events might be said to be internal: (a) the occurrences of the events might be located internally and (b) the characterizations of the events as one or another type of mental event might supervene on what is internally located. It seems to me that Williamson's argument against "the internalist conception of mind" works only because it is directed against (b) rather than (a).

Second, Williamson's identification of the basis of justified belief as everything one knows avoids skepticism, but so does the widely held view that justification starts with all one's present beliefs and experiences. So, the argument from skepticism does not conclusively establish Williamson's identification.

Third and finally, I have suggested that it is more in keeping both with Williamson's goals and with ordinary language to take factive attitudes to be attitudes toward facts rather than propositions and to suppose that factive

attitudes, so understood, are more basic than propositional attitudes.

Finally, let me again emphasize how good this book is and how important it is for contemporary theory of knowledge.

Bibliography

1. Bennett, J., (1976). *Linguistic Behaviour*. Cambridge: Cambridge University Press.
2. Cohen, S., (1984). "Justification and Truth." *Philosophical Studies* 48: 279-95.
3. Dennett, D. C., (1969). *Content and Consciousness*. London: Routledge and Kegan Paul.
4. Dretske, F., (1981). *Knowledge and the Flow of Information*. Cambridge, MA: MIT Press.
5. Feldman, R., (1985). "Reliability and Justification." *The Monist* 68: 159-64.
6. Field, H., (1977). "Probabilistic Semantics," *Journal of Philosophy* 74: 379-409.

7. Fodor, J. A., (1980). "Methodological Solipsism as a Research Strategy in Psychology," *Behavioral and Brain Sciences* 3: 63-73.
8. Foley, R. (1985). "What's Wrong with Reliabilism?" *The Monist*: 188-202.
9. Ginet, C., (1985). "Contra Reliabilism." *Monist* 68, 175-87.
10. Goldman, A. I., (1979). "What is Justified Belief?" In Pappas (ed.), *Justification and Knowledge*. Dordrecht: Reidel, 1-23.
11. Goodman, N., (1955). *Fact, Fiction, and Forecast*. Cambridge, MA: Harvard University Press.
12. Harman, G., (1973). *Thought*. Princeton, NJ: Princeton University Press.
13. Harman, G., (1988). "Wide Functionalism," in Stephen Schiffer and Susan Steele, *Cognition and Representation* (Boulder, Colorado; Westview Press: 1988) pp. 11-20. Reprinted in Harman, G., *Reasoning, Meaning, and Mind* (Oxford: Clarendon Press, 1999), pp. 235-243.
14. King, J., (forthcoming). "Designating Propositions," to appear in *The Philosophical Review*.

15. Loar, B., (1981). *Mind and Meaning*. Cambridge: Cambridge University Press.
16. Lycan, W. G., (1984). *Logical Form in Natural Language*. Cambridge, MA: MIT Press.
17. McDowell, J. H., (1994). *Mind and World*. Cambridge, MA: Harvard University Press.
18. McDowell, J. H., (1995). "Knowledge and the Internal," *Philosophy and Phenomenological Research* 55: 877-93.
19. Moffett, M., (ms). "Knowing Facts and Believing Propositions: The Problem of Doxastic Shift."
20. Parsons, T., (1990). *Events in the Semantics of English: A Study of Subatomic Semantics*. Cambridge, MA: MIT Press.
21. Parsons, T., (1993). "On Denoting Propositions and Facts," *Philosophical Perspectives* 7: 441-460.
22. Pollock, J., (1984). "Reliability and Justified Belief." *Canadian Journal of Philosophy* 14: 103-14.
23. Pollock, J., (1999). "Procedural Epistemology: At the Interface of Philosophy and AI," in John Greco and Ernest Sosa, *Companion to*

Epistemology, London: Blackwell.

24. Putnam, H., (1960). "Minds and Machines." In S. Hook (ed.), *Dimensions of Mind*. New York: NYU Press.
25. Putnam, H., (1964). "Robots: Machines or Artificially Created Life." *Journal of Philosophy* 59: 658-71.
26. Putnam, H., (1967a). "The Mental Life of Some Machines." In H. Castañeda (ed.), *Intentionality, Minds, and Perception*. Detroit: Wayne State University Press.
27. Putnam, H., (1967b). "Psychological Predicates." In W. H. Capitan and D. D. Merrill (eds.), *Art, Mind, and Religion*. Pittsburgh, PA: University of Pittsburgh Press.
28. Quine, W. V., (1960). *Word and Object*. Cambridge, MA: MIT Press.
29. Rawls, J., (1971). *A Theory of Justice*. Cambridge, MA: Harvard University Press.
30. Sellars, W., (1963). "Some Reflections on Language Games." In his *Science, Perception, and Reality*. London: Routledge and Kegan Paul.
31. Stalnaker, R., (1984). *Inquiry*. Cambridge, MA: MIT Press.

32. Stampe, D., (1977). "Toward a Causal Theory of Linguistic Representation," *Midwest Studies in Philosophy* 2: 42-63.
33. Steward, H., (1997). *The Ontology of Mind: Events, Processes, and States*. Oxford: Oxford University Press.
34. Taylor, C., (1964). *The Explanation of Behaviour*. London: Routledge.
35. Vendler, Z., (1972). "On what one knows," in Vendler, *Res Cogitans: An Essay in Rational Psychology*, Ithaca, NY: Cornell University Press, pp. 119.
36. Williamson, T., (1999). "Truthmakers and the converse Barcan formula," *Dialectica* 53: 253-70.