

Forthcoming in *Williamson and His Critics*, ed. P. Greenough and D. Pritchard (OUP).

Timothy Williamson's *Knowledge and Its Limits*¹

Ernest Sosa

Timothy Williamson's *Knowledge and Its Limits* brilliantly interweaves themes from epistemology and philosophy of mind, for a radically new position that brings together two disciplines somewhat distanced in recent decades.² As part of that effort, Williamson argues that knowledge is a mental state, powerfully challenging the widespread assumption that knowledge is mental only by courtesy of the contained belief. The *natural* view, we are told, is that knowledge is a mental state as fully as any propositional attitude.

If the content of a mental state can depend on the external world, so can the attitude to that content. Knowledge is one such attitude. One's knowledge that it is raining depends on the weather; it does not follow that knowing that it is raining is not a mental state. The natural assumption is that sentences of the form 'S knows p' attribute mental states just as sentences of the forms 'S believes p' and 'S desires p' do....(6)³

Believing truly, on the other hand, is not a mental state, and hence not an attitude (except by courtesy of the contained believing). This becomes important for the book's later attempt to characterize knowledge as the most general factive, stative attitude, the most general stative attitude that one can have only to true propositions. If believing truly were a stative attitude, then believing truly would

¹ This paper derives from an APA symposium on *Knowledge and Its Limits*. (Oxford: Oxford University Press, 2000).

² As compared with the time of, say, Sellars's "Empiricism and the Philosophy of Mind," though the work of Davidson, Dretske, Goldman, McDowell, Pollock, and Stich, among others, has sustained the connection all along.

³ Parenthetical references in the text are to *Knowledge and Its Limits*.

be a more general factive stative attitude than knowledge, which would falsify the account of knowledge.

Ideally, a claim that knowledge is a mental state would rest on a well-founded account of what makes a state mental, and as usual Williamson does not disappoint. He gives an account of the mentality of states, by way of a theory of the mentality of concepts, one founded on a distinction between those that are intuitively mental and those that are not.

This note has four parts. The first lays out Williamson's account of mental concepts and mental states, and his characterization of knowledge as the most general factive stative attitude. The second problematizes the account of mental states and the characterization of knowledge, and offers an alternative account of when a state is *purely* mental; according to this account, knowledge really is mental only by courtesy of the contained belief (an internalist intuition opposed to the externalism featured in the book). The third part reflects on possible sources and support for such a conception of the purely mental. The fourth and last part takes up the KK principle, Williamson's *reductio* of it, and the consequences of that for the possibility of reflective knowledge.

1. Williamson's Account

a. Mental concepts and mental states

The concept *mental state* can at least roughly be defined in terms of the concept *mental concept of a state*: a state is mental if and only if there could be a mental concept of that state. (28)

If the concept *C* is the conjunction of the concepts *C*₁, ..., *C*_{*n*}, then *C* is mental if and only if each *C*_{*i*} is mental. (29)

... [The] metaphysical contrast [between kinds of mental states] does not immediately entail the conceptual contrast. Nevertheless, it is hard to see why someone should accept one contrast without accepting the other. If the concept *believes truly* is non-mental, its imagined necessary coextensiveness with a mental concept would be a bizarre metaphysical coincidence. If the concept *know* were a non-mental concept of a mental state, its necessary coextensiveness with a mental concept would be an equally bizarre metaphysical coincidence. (29)

On this basis an important contrast is drawn between knowing and believing truly. Knowing that snow is white qualifies as a mental state or attitude while believing truly that snow is white does not. We can factor believing truly into the belief part and the truth part, but we cannot factor knowing similarly. *S believes truly that snow is white* is tantamount to the conjunction *S believes that snow is white & it is true that snow is white*, the second conjunct of which is non-mental. Accordingly, the *concept* of believing truly that snow is white fails to be mental. As for the *state* of believing truly that snow is white, that will be mental only if some other concept of that state, other than that of believing truly, is mental. But there is no apparent reason to think that there is any such concept.

b. A characterization of propositional knowledge

Factive attitudes are ones we can have only to truths: for example, perceiving that such and such, and remembering that such and such. Some such attitudes are states, “stative” attitudes. Knowing, we are told, is the most general factive stative attitude, the one you must bear to a truth if you bear to that truth any of the others.

So much for what a factive stative attitude is, and for the place of knowledge among such attitudes. Now, a factive stative attitude is characteristically expressed in language by means of a *factive mental state operator* (an FMSO). Syntactically these function like verbs, semantically they are unanalysable: that is to say, no such operator is "... synonymous with any complex expression whose meaning is composed of the meanings of its parts." (34) Any such operator \emptyset must also meet three further conditions:

First, \emptyset typically takes as subject a term for something animate and as object a term consisting of 'that' followed by a sentence. Second, \emptyset is factive, in the sense that the form of inference from 'S \emptyset s that A' to 'A' is deductively valid.... Third, 'S \emptyset s that A' attributes a propositional attitude to S. (34-5)

The ensuing discussion of FMSOs may be summarized as follows (as it is by Williamson on p. 39): '*Know*' is an FMSO, and if \emptyset is an FMSO, then from 'S \emptyset s that A' one may infer both 'A' and 'S knows that A'.

In the material mode, the claim is that knowing is the most general stative propositional attitude such that, for all propositions p, necessarily if one has it to p then p is true. (39)

It is crucial that 'believes truly' *not* count as an FMSO. From 'knows' one may infer 'believes truly' but the converse inference is invalid. If 'believes truly' were an FMSO, therefore, believing truly would be a more general stative propositional attitude than knowing, so knowledge would not be the most general such attitude, and the account would be refuted. Williamson recognizes this threat and blocks it by requiring that FMSOs be semantically unanalysable, a condition that 'believes truly' does not meet.

So I begin by focusing on two main ideas in the book. First is the idea that knowledge is a mental state as fully as belief or any other propositional attitude. Second is the characterization of knowledge as the most general factive stative attitude.

2. A critic's critical comments

a. Safe belief and the account of knowledge as the most general factive stative attitude.

Consider first how to understand *safe* belief, which is a kind of reliable belief. As we are reminded later in the book, the "... argument of Chapters 4 and 5 connected knowledge and safety. If one knows, one could not easily have been wrong in a similar case." (147) According to Chapter 5, "... one avoids false belief reliably in [a given case alpha] if and only if one avoids false belief in every case similar enough to [alpha]." (124) Compare with this the following two conditions:

- i. One would believe *p* only if it were true (or, alternatively, one would not believe *p* without its being true).
- ii. Not easily would one believe *p* without being right.

A belief safe in sense ii can still be false, so long as it is of a relevant sort enough instances (or a sufficient proportion) of which would be true. But Williamson's notion of safe belief requires i, not just ii. It requires that one avoid false belief in *every* case similar enough to the actual case of belief. And there is no case more

similar to the actual case than the actual case. So if a belief is *safe* then it must be true.

That being so, safe belief is itself a factive stative attitude if a stative attitude at all. Williamson hence could not consistently allow that knowing implies safe believing without insisting also that safe believing implies knowing. Once it is allowed that knowing implies safe believing, then only if *safely believes* is another concept of the state of knowledge will the theory stand firm in all its main components. For if safe believing, the state, is entailed by knowing without entailing it, then it is a more general factive stative attitude than knowledge. So knowledge could not then be the most general such attitude.⁴

Based on a requirement of safety, Williamson attacks the luminosity of our mental states, and the myth of the given. Simple safety is dubious as a general requirement for propositional knowledge, however, and

⁴ It might be replied that ‘safely believes that p’ is not an FMSO, since it is semantically analysable. In considering this reply, let us first recall that semantic analysability is distinguished in the book from syntactical complexity. Thus, we are told that ‘She could feel that the bone was broken’ is (roughly) synonymous with ‘She knew by the sense of touch that the bone was broken’. Compare ‘She could hear that the volcano was erupting’ which is said to be synonymous with ‘She knew by the sense of hearing that the volcano was erupting’. These are said nevertheless to be semantically fused, unanalysable. In what respect, however, is ‘... safely believes that p’ different from ‘... knows by the sense of touch that p’ so that the former is semantically analyzable though the latter is not?

It is open to doubt that ‘... safely believes that p’ is semantically analyzable and fails for that reason to be an FMSO. But if it is an FMSO, then safe belief bids fair to be a factive stative attitude, in which case the account of knowledge as the most general such attitude will stand only if the concept *safely believes* is necessarily equivalent to the concept *knows*. Once committed to the necessary equivalence of *knows* with *safely believes*, one might well claim a solution to the Gettier problem, an analysis of knowledge. Not an analysis in any sense that would require analysans and analysandum to be synonyms (or at least the corresponding expressions in the expression of the analysis to be such), but this has not been the objective since the early days of the Gettier project. Rather it would be an analysis through being an interesting necessary biconditional, especially a noncircular one. It would be tempting now to say that knowledge is just equivalent to safe belief, which would save the account of knowledge as the most general factive stative attitude. Unfortunately, safe belief is clearly not equivalent to knowledge, since belief in any necessary truth would seem safe no matter how ill-formed, as is belief in contingent propositions such as the one that you affirm in thinking you have a brain.

Williamson's attack has been rebutted through rejection of simple safety.⁵ Nevertheless, the attack can be mounted through a more defensible requirement of aptness, according to which in order to know that p, one must believe aptly that p, that is to say, the correctness of one's belief that p must be attributable to an epistemic competence exercised through that belief.⁶ A deep problem for the attack on the given remains nonetheless, namely that the concept of luminosity invoked by Williamson is unfairly and implausibly strong. By his lights what is required to make a condition luminous is that one could not possibly be in it without being in a position thereby to know that one is in it. However, a much more reasonable requirement is weaker: namely, that a condition is quasi-luminous only if there is a degree to which one can be in it, such that if one is in it at least to that degree, then one is in a position to know that one is in it (n.b.: "that one is in it," *not*: "that one is in it to that degree"). Of course, 'luminous' is a technical term with a stipulated meaning. Why then is it "much more reasonable" to invoke the broader quasi-luminosity? Because this seems so much the more charitable way of understanding the doctrines of cognitive home and the given that are put in question by reference to luminosity.⁷

That is an initial sketch of the position. We might then approach the question whether knowledge is a mental state as follows:

- a. X is a mental state only if *consciously X'ing* is quasi-luminous.
- b. Consciously knowing is not quasi-luminous

⁵ See R. Neta and G. Rohrbaugh, "Luminosity and the Safety of Knowledge," *Pacific Philosophical Quarterly* (2004).

⁶ I argue extensively for the substitution in *Apt Belief and Reflective Knowledge*, the 2005 Locke Lectures, forthcoming from Oxford University Press (2007).

⁷ Of course, it remains to be seen whether one's states of knowing qualify as quasi-luminous so as to give us an access to them no less privileged than the access we enjoy to our mental states generally.

c. So, knowing is not a mental state.

Consciously knowing is not a quasi-luminous mental state either because it does not have degrees, or because although it does have at least one degree, it does not meet the other requirements specified above.

b. The purely mental.

Closely related to our present discussion is a further question also taken up by Williamson: namely, whether there is some core of purely mental states that excludes knowledge. Here now is my text for the discussion to follow:

What is at stake is much more than whether we apply the word 'mental' to knowing. If we could isolate a core of states which constituted pure mind by being mental in some more thoroughgoing way than knowing is, then the term 'mental' might be extended to knowing as a mere courtesy title. On the conception defended here, there is no such core of mental states exclusive of knowing. If we want to illustrate the nature of mentality, knowing is as good an example as believing.... For similar reasons, other truth-entailing attitudes such as perceiving and remembering that something is the case may also be classified as mental states. Knowing can be understood as the most general of such truth-entailing mental states.

(6)

Williamson thinks it impossible to isolate any such core of purely mental states. His strategy is to view the mentality of a state as requiring that there

could be a mental concept of that state. But a concept is mental only if it is *not* a conjunction of concepts at least one of which is non-mental. Using this basic strategy for conceiving of pure mentality, it is argued persuasively that there is no notion of the purely mental that excludes knowing from that inner core.

However, there is an alternative way of understanding degrees of purity. Token cases of knowing are token cases of believing that satisfy some further conditions.⁸ For example, I argue elsewhere that a belief, i.e., a token believing, qualifies as an instance of animal knowledge only if it is a believing that gets it right because of how it manifests a competence of the believer's. (But the point to follow could be made equally well if we preferred to require truth+justification, or having been reliably formed, etc.) Still, the very token believing that is knowledge might have failed to be knowledge, since it might have been false, for one thing, or might not have been right attributable to any competence of the believer's. The state in question would then seem plausibly to be mental because of its being a believing, in such a way that it would have been mental whether it had been a knowing or not, so long as it had been a believing. That token believing could not possibly have been mental in virtue of being a knowing, however, without being mental in virtue of being a believing. Besides, the mental-making power of its property of being a knowing is plausibly inherited from the mental-making power of its property of being a believing. The token state is, after all, both a believing and a knowing, and it is guaranteed to be a mental state by each of those properties. But it could not have been mental in

⁸ Although the argument to follow is put in terms of token states, I believe that to be inessential. It could be made in terms of mental states directly, where some states, such as belief, are entailed by other states, such as knowing, with the same content.

virtue of being a knowing without being mental in virtue of being a believing, whereas it would have been mental in virtue of being a believing whether it had been a knowing or not. Moreover, it would have been mental whether it had been safe or not, so long as it had been a believing.⁹ So its mentality would seem to derive from its being a believing. Its being a believing that is safe and amounts to knowledge seems inconsequential vis-à-vis whether it is mental or not.¹⁰

That at least is how on the surface it would *seem*, but there is some subtlety beneath that surface. For convenience we are coining the term 'mentalizing' for being in a mental state or engaging in some mental activity. Our question then is whether knowing is subserved by a more thoroughgoing

⁹ Or, to detach from tokens, when one knows one would have mentalized (would have been in a mental state) by being in the entailed belief state, whether one had been in a state of knowing or not, but one would not have been in a mental state of knowing except by being in a mental state of believing with the same content. So one mentalizes by virtue of believing and by virtue of knowing, but one would have mentalized by virtue of so believing whether one had so known or not, whereas one could not have mentalized by virtue of so knowing without mentalizing by virtue of so believing.

¹⁰ An entity's mentality might be viewed as a fundamental, intrinsic, and even essential category of it, as it was by Descartes. Alternatively, the entity might be viewed as fundamentally only physical, and mental in virtue of some contingent, perhaps relational properties of it, as it is by physicalists. If we take the second view, then a state's mentality is not a fundamental property of the state. Mentality, the purity thereof, and the lack thereof, may even concern in the first instance only our concepts. Williamson shows signs of *perhaps* viewing things this way in defining mentality directly for concepts and indirectly for states. A state is said to be mental if there could be a mental concept of it. Now this will in any case depend on what it means for a concept to be *of* a state. One thing it cannot mean is simply that it refers to the state. For the state of the earth of its now quaking is not made mental by being the one that the victims are now thinking about as such a state. What then is it that makes a state one of which a concept is a concept? Perhaps the realizer of a functional concept is a state of which the functional concept is a concept? But then it is hard to set limits on what could possibly be a mental concept. Who knows what possible superphysical entities might realize mental concepts through earthly quaking? And, besides, a single physical state could realize several non-equivalent functional concepts, whereas for Williamson concepts are of the same state only if necessarily coextensive. So I confess to some unclarity on what sort of framework of states and their concepts is used in the book, and in particular on what makes a concept a concept of a certain state.

way of mentalizing, in such a way that to know is to mentalize only by courtesy.

Compare ways of being colored (*chromatically* colored). A surface can be colored by being red, for example, or, alternatively, by being blue. Of course, it can in turn be red by being, more specifically, scarlet. Plausibly, being scarlet is a way, a more specific way, of being colored. And this is so even in the case of something that is both red and scarlet. Despite the fact that this surface would still have been colored even if it had not been scarlet but only red, nevertheless it is colored by being scarlet, and in virtue of being scarlet; and being scarlet is a specific way of being colored.

Compare, further, ways of being shaped. Something can be “shapely” (*regularly* shaped) by being spherical, or, alternatively, by being cubical. Of course, it can also be *surprisingly* spherical. And, necessarily, if surprisingly spherical, then it is spherical. Moreover, it might be spherical without being surprisingly spherical. In some sense, however, being surprisingly spherical is not a more determinate shape property, nor a species of shapeliness. And this remains so even if the property of being surprisingly spherical cannot be factored into a pure shape component, and, as a separate (conjunctive) factor, some non-shape component involving surprise.

I submit that knowing is to mentalizing as being surprisingly spherical is to being shaped, and not as being scarlet is to being colored. Knowing is not a more specific way of mentalizing, just as being surprisingly spherical is not a species of shapeliness. *Knowing* is mental by courtesy of the involved believing analogously to how *surprising sphericity* is a shape state by courtesy of the involved sphericity.

Even someone who shares that intuition might well wonder, however, what more generally lies behind it. What accounts for it? What is it that distinguishes the relation between surprising sphericity and sphericity from the relation between

being scarlet and being colored? One relevant difference is that the surprisingness of the sphericity of a sphere is somehow extrinsic to its sphericity, while the scarletness of the scarlet chromaticity of a surface is not similarly extrinsic to its chromaticity, but is simply a more specific form of the chromaticity.

The proposal is then that knowings can be such *in part* by virtue of being believings that satisfy such further conditions as safety and truth, even if there is no factorization of knowledge into some set of such conditions, where the believing is isolated from all other components. This is similar to how something can be surprisingly spherical in part by virtue of the fact that its sphericity is or would be surprising, even if there is no relevant factorization of being surprisingly spherical into some set of such conditions, one that isolates the sphericity from all other components.

We are thus free to insist that just as being surprisingly spherical is not a more specific form of being shapely (regularly shaped) by comparison with being spherical, so knowing is not a more specific way of being mental, by comparison with believing. This lends plausibility to the notion that knowing is mental if at all then only by courtesy of the involved believing, just as being surprisingly spherical is a regular-shape state if at all then only by courtesy of the involved sphericity.

3. Further reflections

I confess that it is not yet clear to me how much turns on this issue of taxonomy. Why does it matter just where we draw the boundaries of the purely mental? How does the concept of the mental connect with human concerns so as to sustain our intuitive drawing of the boundaries where we do draw them? What connections with what concerns speak in favor of so drawing them?

Williamson adduces much subtle and persuasive reasoning for the view that factive states are explanatorily relevant beyond any contained non-factive states, and for the view that knowledge in particular is a prime example. But why should that tend to show that knowledge is mental, and mental not just by courtesy of the contained belief? It is not as though only seriously mental states can have explanatory efficacy. If he were right, that would of course tend to show that our concepts of factive mental states do connect with important human concerns, those involved in the prediction and explanation of human conduct. What remains unclear is why this should be thought to imply that knowledge, along with other factive stative attitudes, must be mental in the most serious and underivative way. Even less is it clear why, compatibly with its explanatory importance, knowledge cannot be mental only by courtesy of the contained belief.

In addition, other human concerns, besides our interest in the explanation of human conduct, may lead us to distinguish those mental states that are purely mental. Are there such concerns?

Before we go into that question, it bears notice that the distinction between states that are purely mental and those that are not is arguably built into common sense and ordinary language. This would help explain why individuals draw the distinction as they do, but not why the community does so. What wants explaining is why we join together in making the distinction between the mental and the non-mental as we do, and more specifically why our relevant intuitions permit a delimitation of the purely mental as above. Assuming we do have such a conception of the purely mental, one may well wonder about its sources. That it is part of common sense and reflected in ordinary language is not the sort of answer sought, since what we want to understand is why we make the distinction as we do intuitively enough that it counts as part of common sense and ordinary language.

Here I would like to sketch one possible source of our interest in the purely mental: namely, its connection with issues of basic normative responsibility, of proper praise or blame, or at least of proper admiration and contempt. Plausibly it is to the purely mental that we look for basic determinants of such responsibility. This is a possible source of our distinction, or at least a correlative stance. Consider, for example, the victim of Descartes's evil demon, and focus not just on her beliefs and their epistemic standing, but also on her choices, intentions, policies, and character. Detached from physical surroundings, her mental states and episodes, practical and theoretical, could not have their normal consequences, and may be entirely ineffectual physically. Nevertheless, is it not powerfully intuitive that such a victim may be no less admirable or contemptible, no less worthy of praise or blame, than a twin so plugged into its surroundings that its mental goings-on do have their normal physical effects?

This sort of general human concern has a more specific epistemic form, involving our interest in epistemic responsibility, in proper epistemic praise or blame, admiration or contempt. Compatibly with this, our distinction might of course also derive from some less direct important sources. For our interest in the assessment of ourselves and our fellows does itself have proper sources of its own, including the desire for self-improvement and the need for proper allocation of trust. In conclusion, I would like to sketch how the more direct sources specified (and their own sources in turn) comport with our account of the purely mental. Consider then our interest in epistemic responsibility, credit and blame, admiration and contempt. How does this interest bear on the richness of our epistemic conceptual repertoire, and specifically on how this repertoire goes beyond the concepts of truth and reliability?

A belief that constitutes knowledge must be appropriately derived. But it is not enough that the process whence it derives be a reliable process. It must also be

a process appropriately related to the subject, by manifesting a competence or intellectual virtue seated in that subject. Our epistemic conceptual repertoire permits the assessment of a belief as true, of course, and also as “reliable.” But these two concepts would yield only an impoverished epistemology. Unless we go beyond them we would be unable to do justice to the intuitions engaged, for example, by the following two cases.

New Evil Demon. If things appear the same to two subjects from the inside, and if this is so throughout their lives, and if all along they are the same in how they manifest their relevant dispositions to be in purely mental states, then they cannot differ in their respective degrees of justification for any present belief. At least there does seem to be some such internal, or subjective, justification that a belief might have independently of being true or reliable.¹¹

Clairvoyance. A belief might be true and even safe despite the subject’s lack of justification for hosting it. Because Claire has been blessed with a special faculty that puts her reliably in touch with facts that normal humans so situated could not discern, she finds herself believing things quite reliably and correctly, things that she evidently could not be remembering (by construction of the example). So here would be a sort of case where the reliable (and, of course, correct) believing would come detached from justification (in some appropriate, internal sense).

¹¹ This might be thought incompatible with proper externalist restrictions on concept possession, but similar thought experiments escape this objection.

Various proposals are on offer as to how we might best understand such cases. Here are three examples, each of which profiles a kind of justification allowed to the Demon's victim while denied to Claire the clairvoyant. Each specifies a condition that any belief of yours would have to satisfy in order to be justified.

Foley-rationality: that your belief comport on appropriate reflection with your deepest epistemic standards.

Classical foundationalism (BonJour, Fumerton, Conee, Feldman, Lewis): that your belief relate properly to your evidence, conceived of as your basic experiences (and memories). On one way of developing this approach, the belief must result from appropriately taking the given, either through introspection or through perception, and reasoning properly from there, with the strategic aid of memory all along the way.

Virtue contextualism and perspectivism (disclosure: this is an approach that I myself favor): that your belief derive from the good performance of cognitive virtues seated in the subject (features whose operation would generally deliver relevant epistemic goods, such as truth, in the actual world, for beings of your kind in your normal habitat), a fact that does not escape your reflective awareness.¹²

Obviously, these are barely sketches. Still they indicate further epistemic desiderata to which epistemologists have been sensitive. And they all plausibly

¹² For a discussion of whether and how Alvin Goldman also falls under this approach, see my "Goldman's Reliabilism and Virtue Epistemology," in a festschrift volume of *Philosophical Topics* (2002), and Goldman's "Reply," *ibid.*

cater to our interest in self- and fellow-assessment in world-independent respects. Even more than the formalized law, with its multiple practical constraints, common sense is interested not only in successful murders but also in botched attempts, no matter how remotely they may have failed. We are also interested in how well an astronaut performs in simulated flight tests and not just in real life performance. Or take a thermostat in a display room, which may perform well or ill, even when it is controlling the temperature of no space. Similarly, we can perform epistemically well or ill even if detached from any environment in which that performance would have its characteristic, desired outcome: namely, reasonable, safe, and apt belief, i.e., knowledge.

In assessing the performance of the thermostat, in judging it to perform well or ill, we attribute what it does to the device itself as *its* doing. In such assessments of a device, we go by the extent to which the outcome performance derives from its relevant character, from its combination of stable mechanisms (input, processing, and output) whose combined good performance would enable it to work well in securing its proper objectives. (“How are the relevant objectives set, and properly set?” An interesting question, increasingly so as we move from artifacts to animals, humans, groups, and back again.) We *credit (blame)* the device for its good (bad) performance under simulation, and on a certain dimension it is no less to its credit (blame) than if it had been a real-life performance.

The concept of the purely mental jibes with this dimension of our self- and fellow-assessment, of what we care about even in a simulational performance. But wherein lies the value of that kind of assessment of one’s own actions or states, or those of one’s fellows? What’s the point of engaging in such assessments? Our interest in them plausibly derives at least in part from our interest in rating ourselves and others as performers in various dimensions. This interest in turn seems plausibly enough to have a variety of sources. It is worth knowing how good

we already are if we wish to improve. It is good to know our flaws if we aim to remove them, our strengths if we aim to preserve them. We want to know how reliable our fellows are, especially those we join in common endeavors. It is hard to set limits on our need for such knowledge, moreover, or on the practical potential for such need. So, again, we are interested in our flaws, and in our strengths, along various dimensions of potential accomplishment. That way we can better tell whom to trust in what circumstances. On the present account, assessing offline performance is of interest because it bears on assessment of the performer, who is properly credited or blamed through such assessment. And assessment of the performer is of interest for the reasons specified, among others.

If that is roughly right, does it mean we are treating each other as little better than thermostats? Not at all. The proper dictum is *not* that one must never treat others as means. This would be impossible for social beings to obey. Others must not be treated *only* as means, true enough: one's interactions must be sensitive to the proper intrinsic respect that our rational fellows deserve. But this is not in the least endangered by the analogy between our dimensions of assessment of thermostats, and our dimensions of self- and fellow-assessment in epistemic respects. Whether something is a mere tool and is properly so treated is independent of whether its performances can be evaluated under simulation in ways that matter enough to earn it credit and trust.

4. The KK thesis and the status of reflective knowledge

Internalists tend to like the KK thesis, while externalists tend to reject it. For externalists, knowledge is belief that satisfies external conditions of causation, tracking, or reliability. You do need to *be* awake, of course, in order to know by

perception, but you needn't know that you're awake. By contrast, internalists require for perceptual knowledge that one know oneself to be awake and not dreaming. Some internalists back up their intuitions with a general principle: *that really knowing requires you to know that you know*, the KK thesis.

As it stands, the KK thesis leads immediately to vicious regress, but a better version avoids the regress:

KK If one knows that p , and considers whether one does, then one knows that one does.

Williamson argues that this still reduces to absurdity if accepted in its full generality. It reduces to absurdity for magnitudes M such that:

W One knows that: if one knows that x does *not* have M to degree i , then x does not have M to degree $i+1$.¹³

Many magnitudes and measures plausibly fit this bill. Indeed it is difficult to find a magnitude that does not admit a measure under which it plausibly fits our bill. So it seems far from generally true that one knows something only if one knows that one knows it.

Williamson's *reductio* highlights the following form of reasoning:

1. One knows x not to have M to degree 0. (Assumption)
2. One knows 1 (By KK)

¹³ This is not exactly Williamson's formulation, but it is a close relative, and seems plausible enough for present purposes.

3. One knows that if one knows 1, then x does *not* have M to degree 1.
(From principle W, above)
4. One deduces from one's knowledge in 2 and 3 that x does *not* have M to degree 1. (Assumption)
5. One knows x not to have M to degree 1 (By intuitive closure)

By iterating such reasoning *mutatis mutandis* one can derive that one knows x not to have M to degree n, for any particular n no matter how high.

That would reduce to absurdity the KK principle, granted just 1 plus an assumption about our limited powers of discrimination and how that affects our reliability and hence our ability to know, an assumption that underlies the truth of 3 and its like. Again, the assumption takes in general the form of principle W above.

The *reductio* is compelling, which sets a problem for anyone who believes in a bi-level epistemology, with a lower level where only conditions of reliable and rational belief are required, and a higher level that also requires rational awareness of one's reliability. Such reflective knowledge and animal knowledge would seem to differ precisely in that the former requires a KK principle like the following, whereas the latter does not.

KK If one knows that p, and considers whether one does, then one knows that one does.

But this is just what Williamson's *reductio* would reduce to absurdity, which may seem to render incoherent the very idea of reflective, perspectival knowledge, or at least to gut it of all interest. Would not the internalist be committed to the view that if one reflectively knows that p, then one reflectively knows that one so knows that p?

Such a bi-level epistemology, with its animal/reflective distinction, offers a defensible way to meet the severally plausible requirements that seem to clash in the internalism/externalism and the foundationalism/coherentism debates. So it is reassuring to find that its distinctive K+K principle is safe from the otherwise damaging *reductio*.

Also reassuring, with some irony, is the fact that traditional skeptical reasonings can be revived with unreduced plausibility and remain about as initially threatening against a kind of reflective knowledge thus conceived. For example, it will still be a problem to see how one can avoid vicious circularity in ascending from animal knowledge that p to rationally defensible knowledge that one enjoys such knowledge through one's actual complement of faculties or virtues.^{14 15}

¹⁴ "Animal" knowledge is understood here to require rationality or reasonableness in its constitutive belief; it is this that makes bootstrapping seem vicious, not rationally acceptable and hence no source of knowledge, not even of the animal grade. That the viciousness is an illusion, both in the Cartesian Circle and in more recent versions of bi-level epistemology is argued in my "Reflective Knowledge in the Best Circles," *Journal of Philosophy* (1997).

¹⁵ It has been helpful to discuss these issues with John Hawthorne, Brian McLaughlin, David Sosa, and Tim Williamson.