


# KT AND THE DIAMOND OF KNOWLEDGE

*Adrian Heathcote*  
*Department of Traditional and Modern Philosophy*  
*University of Sydney*

AINING AN UNDERSTANDING OF the logic of provability has enabled us to see more clearly what are the limits of provability; it is natural to think that a similar understanding of the logic of knowledge will give us a feel for the (probably considerably rougher) edges of the limits of knowledge. Timothy Williamson thinks so—and in *Knowledge and its Limits* has begun the process of describing that periphery.

That there is an edge at all is, of course, philosophically controversial; it would be denied by anti-realists of a verificationist stripe. However, we accept, since Gödel, that there are true propositions of elementary arithmetic that are unprovable in arithmetic; just so, we should accept—by analogy—that there are true statements that are unknowable. An argument called the *Fitch Argument* tells us that it is so. Williamson has long argued that the Fitch Argument cannot by itself refute anti-realism—because the anti-realist is already committed to the denial of some of the principles of classical logic required to derive the anti-realist conclusion. The point is well made.<sup>1</sup> In *Knowledge and its Limits*, however, Williamson looks at what the Fitch argument tells us if we adhere to classical logic: and that is that there are unknowable truths.

I will return to this issue later in the paper, but first I wish to examine a more fundamental matter: the modal logic that Williamson believes is the correct logic of knowledge.

---

<sup>1</sup>Even if one wants to add that it is too little recognised by anti-realists that they are actually *forced* by the argument to adopt a non-classical logic and that it is not at all clear that this can be done consistently, throughout the metatheory in which the anti-realist position is being enunciated.

In a celebrated 1992 paper Williamson looked at certain modal systems and the disjunction principles that such logics could be said to satisfy. He there isolated those that satisfied, what he called, ‘the Alternative Rule of Disjunction’ and a weaker ‘rule of margins’.<sup>2</sup> These were applied to Sorites type problems as a way of expressing the underlying logic of vague predicates. These same logics, reappear in *Knowledge and its Limits* as a way of expressing what Williamson calls the *anti-luminosity* of knowledge.

Before discussing that matter, however, I want to address a more fundamental question: is knowledge really a modal notion at all?

To suggest that it isn’t may seem rather heretical. Since Von Wright’s work in the 1950’s ‘it is known that. . .’ has been thought to be a modal operator, even if it has not been very clear just what system it might be an operator in.<sup>3</sup> In another 1992 paper, as part of his articulation of the anti-realist’s way out of the Fitch Argument, Williamson himself described two epistemic logics—so surely, one might say, knowledge is now known to have a modal character?

I claim not.

It is of course true—and is agreed by all parties—that the expression *it is known that. . .* possesses a property that resembles the T-property ( $\Box p \rightarrow p$ ) and we could even agree that it has the K-property ( $\Box(p \rightarrow q) \rightarrow (\Box p \rightarrow \Box q)$ )—though this makes a claim about the abilities of knowing subjects that, strictly, seems to go beyond the concept of knowledge itself. The real problem arises when we try to define the conjugate of knowledge: epistemic possibility.

When it is the case that *I do not know that p is false*, this may seem to be adequately represented by the  $\Diamond$  of knowledge. But whenever I don’t know that *p* is false I also don’t know that it is true. Our ignorance is thus equally well represented, and expressed, by  $\neg K\neg p$  and  $\neg Kp$ . If one is true the other will be true as well.

When we force the structure of a modal logic onto knowledge we make a distinction where none exists, for we will choose  $\neg K\neg p$  to stand for  $\Diamond p$  and treat  $\neg Kp$  as distinct from it. This is obviously completely different to the situation that exists for alethic necessity: there,  $\neg\Box\neg p$  says something quite different to  $\neg\Box p$ . And when *p* is a necessary falsehood, say *A* &  $\neg A$ , then  $\neg\Box\neg p$  will be false while  $\neg\Box p$  is true. Similarly, if *p* is a necessary truth then  $\neg\Box\neg p$  will be true while  $\neg\Box p$  is false. We can

---

<sup>2</sup>“An Alternative Rule of Disjunction in Modal Logic” *The Notre Dame Journal of Formal Logic*, 33, (1992) pp. 89–100.

<sup>3</sup>Von Wright sketched one such system in his *An Essay in Modal Logic* (Amsterdam: Nth Holland, 1951). However in his *Logical Studies* (London: Routledge, 1957), he discussed (but did not develop) an alternative account that differed on epistemic possibility.

summarize this situation by noting that

$$(1) \quad \neg K\neg p \rightarrow \neg Kp$$

ought to be true in any logic of knowledge, whereas the corresponding statement for necessity ought to be false. But, because of this, the D-principle ( $\Box p \rightarrow \Diamond p$ ) leads to absurdity for knowledge for we have from D and (1)

$$(2) \quad Kp \rightarrow \neg Kp$$

which entails  $\neg Kp$ . Our modal logic will then entail universal scepticism. Since the modal logic KT extends KD, D will be true in KT. The only way out, I suggest, is to deny that  $\Diamond$  is definable for knowledge. Thus, despite the fact that we can define analogs of K and T for knowledge we cannot and should not suppose that there is a modal logic of same.

A diagnosis might help reinforce and clarify the point.  $\neg K\neg p$  and  $\neg Kp$  are equivalent despite the fact that their negations are distinct. For when we negate the first we get  $K\neg p$  which is *not* equivalent to  $Kp$ . Thus the contrapositive of (1) is false and obviously so. Thus when we move back and forth between positive and negative propositions we are in danger of lapsing into confusion, or scepticism, or both. But even if we are careful it remains true that the structure of the basic knowledge propositions is not that of a modal square of conjugates, but rather a triangle of propositions. Moreover, this triangle is what one might expect if ‘knows’ was not a modal operator at all but rather a simple predicate. (Think of *knows* as a predicate applying to true propositions, some of which are positive and the rest negated; the denial simply takes one outside the set to propositions which may be either true or false.)<sup>4</sup>

The upshot is that there is nothing that corresponds to ‘epistemic possibility’ and the idea that *it is known that...* has the form of a modal operator may be due to superficial linguistic form and not deep logical structure.

But perhaps, it might be suggested, we can simply *impose* the structure of a modal logic by interpreting knowledge as the  $\Box$  and leaving off defining epistemic possibility, except as that which is given by the conjugacy relation  $\neg\Box\neg p$ .

I think this kind of ‘shoehorning’ is fraught with danger, and likely to take us further and further away from the concept that we are trying to understand—but in what follows I will assume that something like it can be done—if for no other

---

<sup>4</sup>A second way of seeing what goes wrong with negated knowledge claims is to note that  $\neg K\neg p$  and  $\neg Kp$  can be read under a ‘knows that’ or a ‘knows whether’ interpretation. The ‘whether interpretation’ has the same pattern here as ‘knows that’ and it is tempting to see us as simply dropping into the ‘whether interpretation’ for these two propositions. But when we consider  $K\neg p$  and  $Kp$  the ‘whether interpretation’ and the ‘that interpretation’ are quite distinct, because these two propositions are equivalent under the former but distinct under the latter.

reason than that it will allow us to follow, sympathetically, Williamson’s project, and make criticisms from within. However, toward the end, I will return to these more foundational criticisms and argue that they present an insurmountable obstacle.

So I will assume, for the sake of argument, that knowledge can be interpreted as a modal operator satisfying at least the **K** and the **T** axioms. (That these axioms are satisfied need not be merely for the sake of argument.) Moreover it is but a short step from this to acknowledging that knowledge should obey, what Williamson calls, the Lemmon-Scott rule of disjunction.

(3) if  $\vdash \Box p_1 \vee \dots \vee \Box p_n$  then  $\vdash p_i$  for some  $i$  ( $1 \leq i \leq n$ )

This principle seems to go little beyond  $Kp \rightarrow p$  and in fact, as Williamson notes, holds for **KT** systems in general.<sup>5</sup> And though, as I’ve been stressing, we cannot carry across results that hold for modal systems to our concept of knowledge, we can still recognise those principles that do in fact hold.

With these general considerations behind us, we now consider Williamson’s arguments for anti-luminosity and against knowing that one knows.

## II

A condition  $C$  is *luminous* if a person in that condition, and in a position to know, would in fact know that  $C$  is the case. But, Williamson argues, if knowing that condition  $C$  obtains at some time requires (for reliability reasons) that it obtain at  $t+\epsilon$ , then condition  $C$  must hold at all later times. Which is absurd, therefore no condition can be luminous.

This argument makes the reliability premise bear an awful amount of the load. Can it in fact bear so much? Here is Williamson’s argument for the premise.

Now at  $t_{i+1}$  one is almost equally confident that one feels cold, by the description of the case. So if one does not feel cold at  $t_{i+1}$  then one’s confidence at  $t_i$  that one feels cold is not reliably based, for one’s almost equal confidence on a similar basis a millisecond earlier that one felt cold was mistaken. In picturesque terms, that large proportion of one’s confidence at  $t_i$  that one still has at  $t_{i+1}$  is misplaced. Even if one’s confidence at  $t_i$  was just enough to count as belief, while one’s confidence at  $t_{i+1}$  falls just short of belief, what constituted the belief at  $t_i$  was largely misplaced confidence; the belief fell short of knowledge. One’s confidence at  $t_i$  was reliably based in the way required for knowledge only if one feels cold at  $t_{i+1}$ .

---

<sup>5</sup>We will return to the issue of ‘theoremhood’ at the end of the paper.

But why is it necessary for me to feel cold in the future for me to know that I feel cold now? Surely, my feeling cold now could lead me to reliably know that I feel cold now, despite the fact that, in a millisecond's time, I die suddenly. My future sudden death cannot act backwards to prevent me from having known something a millisecond before.

If we are determined to make reliability a function of endurance over time it surely makes more sense to look at the cold I felt *before*  $t_i$  rather than the cold I feel just after. The cold that I felt at  $t_{i-1}$  will serve equally well, or better, as basis for a reliability claim. In fact if we replace Williamson's crucial premise with the following

$1'_i$ : If in case  $\alpha_i$  one knows that one feels cold, then either in  $\alpha_{i+1}$  or in  $\alpha_{i-1}$  one feels cold

then Williamson's *reductio* no longer goes through. (I have made the alteration to Williamson's crucial premise  $1_i$  as minimal as possible so that it is still consistent with Williamson's argument in the above quotation. Therefore, I maintain, if Williamson's argument were correct the reliability condition would be equally stable as my  $1'_i$  as his  $1_i$ .)

To see further why it is wrong to think of reliability as requiring the stability of some condition into the future, consider the following case. Suppose I see a tree under dim light—not so dim that I cannot at present recognise it as a tree, but such that if the light were to be dimmer by some finite but small amount, I would no longer be able to recognise it as such. There are two scenarios: one, I see it at dawn, the other that I see it at dusk. If Williamson were right and it were the perturbation into the future that mattered, at dusk I don't know that it is a tree—despite the fact that I've been staring at it all day—whereas at dawn I do know that it is a tree, just because the light is increasing and I will see it more clearly ten minutes from now. This is absurd.

Note also that requiring stability into the future would contradict Williamson's idea that knowledge is a state of mind. For how could my state of mind at  $t_i$  be *constituted* by what will happen in the future? It is plausible to think that a state of mind could be determined by past circumstances, but not by future circumstances. (In fact, the idea that knowledge is dependent on future conditions, combined with the doctrine that knowledge is a state of mind, would surely mean that nothing could be known.)

But this goes to a more general problem: the more we make knowledge conditional on the obtaining of circumstances that go beyond the physical state of the organism and its immediate environment, the more the notion of knowledge being a state of mind is evacuated of sense. (But I will return to the notion of knowledge as a state of mind at the end of the paper.)

So should we make reliability a matter of the endurance of some condition over time? It is not obvious that we should.

What reliability requires is high correlation: we want there to be a high correlation between the obtaining of some condition and my knowing that it is so (assuming that I am suitably attentive, *etc*). What Williamson does is assimilate reliability to the notion of *stability*: so that to know I must have knowledge under perturbations of the conditions. But these are not obviously the same. My computer may reliably crash when I press a particular combination of keys, but not under some perturbation of that action. Moreover stability (in the dynamical systems sense in which Williamson uses the idea) is not obviously an integral part of the concept of knowledge at all.

Thus suppose that in a particular region of France the superstitious farmers never plant cypresses in prime numbers. Alice knows this and looks at a field: she knows that there aren't 17 trees there, but doesn't know that there aren't 14, 15, 16, or 18. Her knowledge is thus not stable—but it is, we may suppose, reliable.

The point is that Williamson's argument should not make us pessimistic that we can formulate a notion of reliability for knowledge. And when we do so it is likely that there will be perfectly legitimate cases of luminous conditions. And the state of my being freezing cold at dawn will surely be an obvious candidate.

If Williamson's case against luminosity fails—as I've argued it does—then it has ramifications for many of the conclusions in the book. I now look at a variant of the luminosity argument that is directed against the so-called *KK*-thesis.

### III

Person A looks at a tree but cannot tell how tall how it is. Williamson wants to represent their self-knowledge of their own ignorance by the following claim: person A knows that if the tree is  $i+1$  inches tall then they don't know that it isn't  $i$  inches tall. (Call this  $1_i^*$ .) The conditional that A knows is permitted to be a weak material implication. Let  $K\alpha$  stand for the claim that person A knows that the tree is  $\alpha$  inches tall.

Combined with a plausible, and only very slightly restricted, version of the *KK*-thesis ( $Kp \rightarrow KKp$ ) and other harmless principles, this leads to the following conditional

$$(4) \quad K\neg i \rightarrow K\neg(i + 1)$$

This again is the main engine in an inductive argument to the conclusion  $K\neg 666$ —which is false since the tree is in fact 666 inches tall.

There is no question that the argument to the conclusion is valid (though in the text it is overly complicated; it can be reduced to a five line conditional proof).

Williamson thinks that the premise that should be rejected is the  $KK$ -thesis. I disagree. In my view the premise that should be rejected is  $1_i^*$ . Moreover it shares many of the weaknesses that led us to reject its namesake in the anti-luminosity chapter.

The first thing to notice about  $1_i^*$  is that it is expressed in such a way that its meaning is hard to grasp. But we can turn it around, contrapositively, so that it is much plainer.

$$1_i^*: K(K\neg i \rightarrow \neg(i+1))$$

But once plain it seems plainly false; for it says that A knows that if he knows that the tree is not  $i$  inches tall then it is not  $i+1$  inches tall. The considerations about the reliability of A's ignorance certainly do not seem to warrant us attributing so much knowledge to A. In virtue of being short-sighted A seems to know altogether too much.

But we do not have to rely on intuitions about A's seeming falsity. For we can combine it with obvious truths to produce obvious disasters. By the facticity of knowledge  $1_i^*$  implies

$$(5) \quad K\neg i \rightarrow \neg(i+1).$$

But we also have, by facticity,

$$(6) \quad K(i+1) \rightarrow i+1.$$

But now the contrapositive of (5), combined with (6), leads to

$$(7) \quad K(i+1) \rightarrow \neg K\neg i.$$

But this means that if A knows that the tree is 666 inches tall then he does not know that it isn't 665 inches tall. This is false, consequently  $1_i^*$  is false.

In my view we should not be too surprised to find  $1_i^*$  false. After all, it fits very ill in an argument whose purpose it is to undermine the  $KK$ -thesis. This is because it seems to be much stronger than the  $KK$ -thesis itself; it allows A to come to know that if he knows that the tree is not  $i$  inches tall then it is not  $i+1$  inches tall either. This last statement amounts to something even more than a  $K\neg K$ -thesis ( $\neg Kp \rightarrow K\neg Kp$ )! Because of this it seems unreasonable to employ  $1_i^*$  in an argument against the  $KK$ -thesis. If, simply by reflecting on his situation, A can come to know  $1_i^*$  then surely he can come to know that he knows that  $p$  simply by reflecting on the fact that he knows  $p$ . If Williamson wishes to maintain  $1_i^*$  then he must find a stronger argument for it than the one in the text—one that does not seem to buy back the very luminosity that he is concerned to deny.

Note also that if, instead of  $1_i^*$  we had started with an uncontroversial truth about A's knowledge—something that follows from its facticity—

$$(8) \quad K(K(i) \rightarrow \neg(i+1))$$

then the same argument that Williamson employs, using the  $KK$ -thesis in the same way, would take us to the harmless truth  $K(i) \rightarrow K\neg(i+1)$ . But if we now combine Williamson’s conclusion with this then we have

$$(9) \quad Ki \vee K\neg i \rightarrow K\neg(i+1).$$

But in fact it seems implausible that A should have knowledge of a conditional that in any way resembles  $1_i^*$ —even when the conditional is a weak material implication. It is plausible to suppose that A does not know that the tree is  $i$  inches tall, and—granting some luminosity—that he should even know that he doesn’t know this. So we can perhaps grant

$$\dots K\neg K(i-1), K\neg K\bar{i}, K\neg K(i+1), \dots$$

over some substantial range of  $i$ . But we should be wary of being led to change what A knows to some conditional statement inside the scope of the first of these operators. A is short-sighted and knows he is—but he has not gained some compensatory *knowledge* of what implies what.

However if A’s knowledge is restricted to this kind of categorical knowledge of his own ignorance then no modification of Williamson’s argument is likely to cast doubt on the  $KK$ -thesis. Nor in my view should it—for the vast number of instances of the  $KK$ -thesis are likely to be true.

#### IV

Given the failure of the arguments against luminosity and the  $KK$ -thesis we should suspect that there is some subtle conflict between the facticity of knowledge and the margin principles that are being utilised in these two arguments: the reliability constraints and the rule of margins principle—for the former is really just a margin of success, as opposed to a margin of error, rule. In the anti-luminosity argument the reliability of knowledge acts as a ‘lock’ to prevent knowledge from smoothly changing into non-knowledge as conditions change; in the anti- $KK$  argument it acts to prevent knowledge of what is not changing to knowledge of what is. Reliability and the rule of margins are used by Williamson to drive knowledge out of the mainstream cases.

But as suggested earlier, the Lemmon-Scott rule of disjunction is a very plausible principle for an account of knowledge to satisfy, since it really just generalises the facticity of knowledge. Thus any principle that is in conflict with it must suffer the gravest suspicions. But the rule of margins is in subtle conflict with it as indicated by the following remark from Williamson’s Appendix 2: ‘KTB is exactly the logic for knowledge determined by the simplest version of the margin for error considerations.’

Why does this indicate a conflict? Because the B (for “Brouwersche”) axiom is not an axiom that any concept of knowledge *should* satisfy. It says

B:  $p \rightarrow K\neg K\neg p$

This is false: it is not the case that just if a proposition is true then I know that I don’t know that it is not true. This fails for all those propositions that are true but not known to be true, let alone those that are true but unknowable. Moreover the system KTB does not provide the Lemmon-Scott rule of disjunction.

This means that either (*pace* Williamson’s remark, above) we should not be interested in the *simplest* version of the margin for error considerations, or we should not be interested in the margin for error rule at all. This is because, dropping B, KT *does* support the Lemmon-Scott rule of disjunction and also the margin for error rule. However if we added the *KK*-thesis to KT we would have KT4 and this has the Lemmon-Scott rule but not the margin for error rule.

But should we even think that KT is the correct modal logic of knowledge? A remark of Williamson’s casts doubt even on that. He says,

A feature of both KT and KTB is that, for any formula A,  $A \supset KA$  is a theorem if and only if either  $\sim A$  is a theorem or A is (Williamson 1992a). This is a formal analogue of the hypothesis in Chapter 4 that only trivial conditions are luminous. (p. 306)

But, once again, this seems quite wrong: the condition  $A \rightarrow KA$  says something far stronger than that trivial conditions are luminous, it says that all provable and refutable statements are known to be true. In the 1992 paper the concept under consideration was not knowledge but clarity (logical truths and logical falsehoods are clearly true and false, respectively)—and here there is some plausibility. But when applied to knowledge it seems false and out of place.

These considerations should make us reflect again on the purpose of seeking formal models for ordinary concepts like knowledge. Have we really made any gain by interpreting knowledge as a modal operator? Or have we been seduced by the idea that progress can be made when we go to formal models—even when it is not clearly progress toward the truth?

When C. I. Lewis extended the propositional calculus by adding modal axioms it was clear what he was trying to do: he was wanting to capture the idea that logical truths were necessary and even necessarily necessary. It was appropriate to extend an axiomatic system because we were interested in theoremhood in the system. Likewise we can make sense of provability in a formal system. But with knowledge we are not extending an axiomatic system, and it makes no sense to think of it as a comparable enterprise. Theoremhood is irrelevant to our purposes.

Rather than theorems and axioms what we are presented with in epistemology is a set of statements, some of which are known. We want to be able to draw inferences from the set that is known—but not much is possible other than that those propositions that are known are true, by the facticity of knowledge.

It is a perennial temptation to make progress on a problem by reducing it to another for which there is a known solution. The imposition of a modal structure onto the concept of knowledge is, I maintain, just such a move. But the superficial resemblance that knowledge bears to necessity is undermined by the fact that absence of knowledge does not at all resemble the absence of necessity: as I argued in section I, there is no diamond of knowledge.

But once the imposition is made the very fact that progress is possible can look like a vindication of the method.

From this point it is tempting to either ‘go normative’ or ‘go naturalist’—adding other constraints and inference rules that either say what we *should* know if we know  $p$ , or what the cost of knowing  $p$  might be, with the kind of limitations that creatures like us have. The *KK*-thesis is in the normative fork, and the reliability and margin for error principles are in the naturalist fork. But neither are really constitutive of the concept of knowledge itself. The concept of knowledge, for example, does not dictate that there should be a lot of nearby conditions that would have given rise to the same knowledge had they obtained.

## V

I end with a few remarks on matters that do not depend on interpreting *knows that* in modal logic.

Williamson believes that knowledge is a state of mind, though believing truly is not. This is an intriguing thesis, but one that would require an argument whose weight is equal to the thesis’ counter-intuitiveness. However the argument when it comes is disappointing, for it amounts to little more than the claim that there is a way of describing knowing and believing truly so that the former is *classified* as a state of mind whereas the latter is not.

The basis Williamson offers is *syntactic*: believing truly is conjunctive—with one conjunct non-mental—whereas knowing is not conjunctive at all. Having a non-mental conjunct is a sufficient condition for a concept to be non-mental.

This is implausible: ‘knowing  $p$  and  $p$  is true’ is conjunctive and logically equivalent to knowing  $p$ . If knowledge *were* a state of mind this concept would surely be too. If it is not, then, since it does nothing more than pull out one of the features of knowledge and restate it as a conjunct, knowledge is not either.

Williamson’s criterion is only going to be persuasive if there is no better basis for a taxonomy than the one he has suggested. But there is. Knowledge is non-mental,

just as believing truly is non-mental, because it is factive. If my knowing that the battle of Agincourt occurred in 1415 depends on it being true that it occurred in 1415, then it is plausible to say that it is not a state of mind: no amount of inner searching will reveal that I know it. Factive states are not in the head.

But how, one wonders, does all of this interact with the idea that there are unknowable truths? Particularly, when we can easily see that these unknowable truths are true. Consider, for example, the statement *the diagonalization of  $x$  is unknowable*. The diagonalization of this statement is

D: The diagonalization of *the diagonalization of  $x$  is unknowable* is unknowable.

If D is false then it is unknowable (by facticity); if D is true then it is unknowable. Therefore the statement is unknowable. And so it must be true. But what, then, is stopping me from knowing it?

The limits of knowledge are strange indeed.

\*

For all the above disagreements I hope I have succeeded in conveying the sense that Williamson's work is an exciting development in epistemology—perhaps the most exciting of the last fifty years. There is much in it that will define how the subject is thought about for a long time to come.