Matthias Steup

Are Mental States Luminous?

Against Luminosity

In *Knowledge and its Limits*, Timothy Williamson gives a fascinating argument against the claim that mental states or conditions are luminous. He defines luminosity as follows:

> A condition C is luminous iff for every case α, if in α C obtains, then in α one is in a position to know that C obtains.[1]

Suppose headaches are luminous. If so, then whenever one has a headache "no obstacle blocks the path to knowing" that one has a headache.[2] To acquire knowledge of one's headache, one merely needs to consider whether one has a headache. Obviously, non-mental conditions are not luminous. Consider the presence of milk in my refrigerator. My refrigerator's having milk in it is not a luminous condition since I'm not *always* in a position to know whether it obtains. For example, right now I have forgotten whether there's milk in the fridge. Hence I'm not right now in a position to know whether there is milk in the fridge. I'd have to go and look.

To show that mental conditions are not luminous, Williamson considers a particular example of a mental condition, that of *feeling cold*, and argues that it is not luminous. From this he generalizes: a parallel argument can be made for any other mental condition. Hence we get the result that no mental conditions are luminous.

For the purpose of deriving the conclusion that feeling cold is not luminous, Williamson describes a case, subsequently referred to as case α, in which "one feels cold at dawn, very slowly warms up, and feels hot by noon."[3] The case is divided into a series of times $t_0, t_1, \ldots, t_n$ leading in one-millisecond intervals from dawn to noon. So α_0 is the case at dawn, α_n is the case at noon, and α_i is the case at a time t_i where $0 \leq i \leq n$. The noteworthy features of the case are the following:[4]

---

[1] Williamson 2000, p. 95.
[2] Ibid.
[3] Ibid, p. 94.
[4] Ibid.

(i) In α_0 one feels cold.

(ii) In α_n one does not feel cold.

(iii) One's feelings of heat and cold change so slowly during this process that one is not aware of any change in them over one millisecond. (NAC, for no awareness of change.[5])

(iv) Throughout the process one thoroughly considers how cold or hot one feels.

Features (i) through (iv) are meant to describe an ordinary case in which nothing controversial is going on. The argument based on this case is of the *reductio* type. Williamson asks us to assume that feeling cold is a luminous condition. Given assumption (iv), if feeling cold is luminous, then at any given time during the dawn-noon interval at which one feels cold, one knows one feels cold. Thus we have:

(L) If (cold) in α_i, then K(cold) in α_i.[6]

In addition to (L), Williamson assumes a second key premise, namely

(R) If K(cold) in α_i, then (cold) in α_{i+1}.

Given (L) and (R), we can derive that one feels cold in α_n. This contradicts assumption (ii). Hence we must conclude that either (L) or (R) is false. Williamson claims that we must select (L) as the culprit since (R) is backed up by solid reasons. Let's first see how the conjunction of (L) and (R) produces the contradictory outcome that one believes one is cold in α_n. Then we'll have a look at Williamson's argument for (R).

The contradiction results as follows: Given assumption (i) and (L), we have:

(1) K(cold) in α_0.

Given (1) and (R), we have:

(2) (cold) in α_1.

Given (2) and (L), we have:

(3) K(cold) in α_1.

Given (3) and (R), we have

(4) (cold) in α_2,

[5] Here I'm following Ramachandran 2005.

[6] Let '(cold)' stand for 'one feels cold' and 'K(cold)' for 'one knows one feels cold'.

and so forth. Eventually, continued applications of (L) and (R) yield the unbecoming outcome:

(cold) in α_n.

As already mentioned, the unbecoming outcome gives us a reason to reject (L) only if (R) is well defended. What, then, is Williamson's defense of (R)?

Williamson motivates (R) by appeal to the plausible thought that knowledge requires reliability. According to Williamson, this requirement is to be understood as follows:

(P_R1) If one knows that *p* in a given case, then *p* is true in every similar case in which one believes that *p*.

Now suppose we have:

(a) K(cold) in α_i.

Since knowledge requires belief, we also have:[7]

(b) B(cold) in α_i.

What is going on one millisecond later in α_{i+1}? NAC tells us that, even though in α_{i+1} one feels slightly less cold, one is not aware of this change. Hence, for any interval α_i - α_{i+1} such that B(cold) in α_i, one will in α_{i+1} believe one is cold. So we have:

(c) B(cold) in α_{i+1}.

The step from (a) to (c) supplies Williamson with another premise:

(P_R2) If K(cold) in α_i then B(cold) in α_{i+1}.

Since cases α_i and α_{i+1} are only one millisecond apart, and since one only feels only slightly less cold in α_{i+1} than one did in α_i, cases α_i and α_{i+1} are very similar. Given that the two cases are similar, (P_R1) and (P_R2) yield (R).

NAC: The No-Awareness-of-Change Assumption

How can the luminosity of the mental be defended against Williamson's argument? Since the only substantive premises are (L) and (R), a defense of (L) would have to aim at finding a flaw in Williamson's defense of (R). Some have argued that, in defending (R), Williamson has made questionable assumptions about the nature of knowledge.[8] I will

[7] Williamson thinks of knowledge requiring confidence rising to the level of belief. In the present context, this terminological difference is irrelevant.

[8] See Brueckner and Fiocco 2002, and Neta and Rohrbaugh 2004.

argue that Williamson's defense of (R) rests on something else there is good reason to view as problematic, namely NAC: the assumption that, in case α, there are one-millisecond changes of feeling less cold than before of which one is not aware. Williamson articulates this assumption as follows:

> NAC Suppose that one's feelings of heat and cold change so slowly during this process that one is not aware of any change in them over one millisecond.[9]

During the dawn-noon interval, one continuously undergoes changes of feeling less cold than a moment before. Williamson assumes that, for one millisecond intervals, one is not aware of these changes. There is a weak and a strong sense in which one can fail to be aware of something. I can fail to be aware of something that's right before my eyes simply because due to inattention I did not notice it. For example, there might be an empty coffee cup on my desk without my being aware of it. Although I'm not aware of it, I'm nevertheless in a position to know that there is an empty cup on my desk. All I need to do is notice it. This example illustrates the weak sense in which I might not be aware of something. In the strong sense, one fails to be aware of something because cognitive access to it is impossible.

When Williamson assumes that in case α there are unnoticed episodes of feeling less cold, he has the stronger sense in mind: in case α, one is not aware of one-millisecond changes because, since they are so short, they are *not noticeable at all*. Lasting only one millisecond, these changes of feeling less cold than before are so tiny that it does not seem to one that one feels less cold than before. In general terms, the point is that, as Williamson himself puts it, "it is metaphysically possible for experience to appear other than it really is."[10] So according to Williamson, each time in case α a one-millisecond change of feeling less cold than before occurs, one experiences reality other than it really is: it appears to one that no change of feeling less cold has occurred when such a change has in fact occurred.

[9] Williamson 2000, p. 94.
[10] Williamson 2005, p. 474.

Indiscernible Episodes of Feeling Less Φ

The following diagram illustrates, for a mental condition of feeling ϕ, the conception of experience that Williamson assumes:

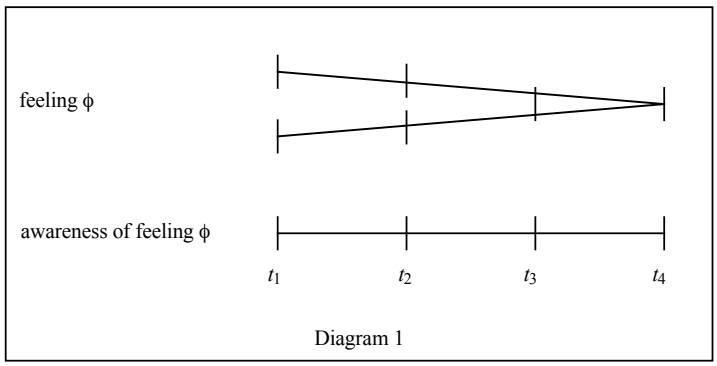

Diagram 1

Consider first one's feeling ϕ.The degree to which one feels ϕ is represented by the distance between the two converging lines. At t_2 the distance between the lines is less than at t_1; so at t_2, one feels less ϕ than one did at t_1. Towards t_4, the lines continually converge; approaching t_4 one feels continually less and less cold. Finally, at t_4, the lines meet; one feels ϕ no longer. Next, consider one's awareness of feeling ϕ. The three intervals t_1-t_2, t_2-t_3, t_3-t_4 represent the smallest possible intervals of noticing a difference in the degree of one's feeling ϕ. Let us call such intervals *minimally discernible* changes, as opposed to *indiscernible* changes. Even though during the t_1-t_2 interval one continues to feel less ϕ, one is not aware of feeling less ϕ until one is at t_2; only then is one aware of feeling less ϕ than one did before. Likewise, one will be aware of feeling less ϕ again only at t_3 and t_4, even though, throughout the t_2-t_4 interval, one has *continuously* felt less ϕ.

So in between each minimally discernible change of feeling ϕ, one undergoes a multitude of *indiscernible* changes of feeling ϕ. That's how we must understand Williamson's key assumption about case α. In case α, we have many intervals of the following kind: In α_{i+1} one feels less cold than one did in α_i, but this change is indiscernible. Therefore, one remains unaware of it. If we assume that a minimally *discernible* change takes at least one second, then we have within that second many *indiscernible* occurrences of feeling less cold than before.

Lucy the Luminosity Friend

Next meet Lucy, an advocate of the luminosity of feeling cold in particular and mental states in general. Lucy holds that, according to her conception of what it is to *feel* ϕ, there is no such thing as feeling ϕ without being aware of it. If you are not aware of feeling an itch, then you don't feel an itch. If you are not aware of feeling pain, then you don't feel pain. And if you are not aware of feeling cold, then you don't feel cold. The entailment, according to Lucy, also holds in the opposite direction. If you are aware of feeling pain, then you feel pain, and if you are aware of feeling cold, then you feel cold. In general terms, Lucy holds that there is no feeling ϕ without being aware of feeling ϕ, and no being aware of feeling ϕ without feeling ϕ. In short, feeling ϕ ↔ being aware of feeling ϕ, or Fϕ ↔ A(Fϕ).

What holds for feeling ϕ, Lucy will say, also holds for *feeling less* ϕ *than a moment before*.[11] If you are not aware of feeling less cold than a moment before, then you do not feel less cold than a moment before. And if you are aware of feeling less cold than before, then you feel less cold than before. The same applies to other mental changes such as feeling less tired than before, feeling less pain than before, or feeling less hungry than before. In general, feeling less ϕ ↔ being aware of feeling less ϕ, or $F_L\phi \leftrightarrow A(F_L\phi)$.

Given that Lucy holds $F_L\phi \leftrightarrow A(F_L\phi)$, she will say that the phenomenology of feeling less cold, as illustrated by Diagram 1, gets things wrong. If in between the times t_1 and t_2 there are no further times at which one is aware of feeling less cold than before, then after t_1 and before t_2 one's feelings of cold and heat remain unchanged between t_1 and t_2. So according to Lucy, the correct phenomenology of feeling less cold is illustrated by Diagram 2.

[11] I will use 'feeling less ϕ' as a short for 'feeling less ϕ than a moment before'.

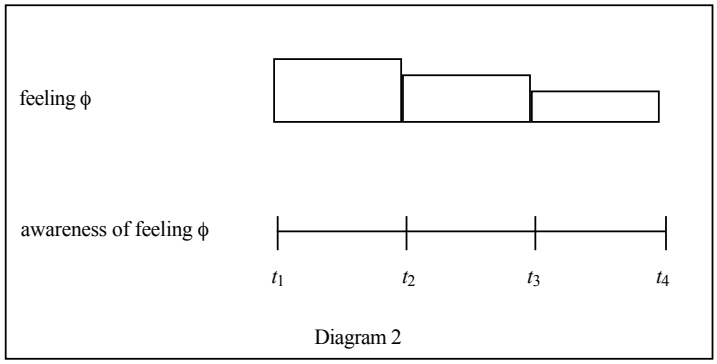

Diagram 2

Diagram 2 differs from Diagram 1 in the following respect: it does not show any indiscernible occurrences of feeling less ϕ. Assume that at t_1, one feels ϕ to a clearly discernible degree. At t_4, one has stopped feeling ϕ. During the t_1-t_4 interval, one's awareness of feeling ϕ changes twice: at t_2 and t_3. One's feeling ϕ also changes twice, namely at the very same times. In between these times, one's awareness of feeling ϕ does not change. Neither does one's feeling ϕ. So one's feeling ϕ and one's awareness of feeling ϕ change concurrently. A change in one does not occur without a change in the other. So according to Diagram 2, the smallest changes of feeling ϕ coincide with the minimally discernible changes of feeling ϕ. According to Lucy, that these changes so coincide is a matter of metaphysical necessity. Feeling cold works like that, and so does any other phenomenal mental condition, such as feeling tired, feeling hungry, or feeling pain. So according to Lucy, when it comes to mental conditions of the feeling ϕ kind, there is no such thing as an indiscernible change. If one feels cold at t_1, and one does not notice any difference in the degree to which one feels cold between t_1 and t_2, then in between t_1 and t_2 the degree to which one feels cold remains unchanged.

Given that Lucy holds $F\phi \leftrightarrow A(F\phi)$ and $F_L\phi \leftrightarrow A(F_L\phi)$, what will she say in response to Williamson's anti-luminosity argument? Obviously, Lucy will say that case α, since it involves episodes of feeling less cold in which one is not aware of feeling less cold, is metaphysically impossible. Lucy will not think, therefore, that Williamson has supplied her with a successful argument against the luminosity of feeling cold in particular and the luminosity of phenomenal mental states in general.

Two Readings of NAC

Let us again have a look at the exact wording of the passage in which Williamson articulates NAC:

> Suppose that one's feelings of heat and cold change so slowly during this process that one is not aware of any change in them over one millisecond.

There are two things about this passage that I find problematic. First, the passage is not ideally explicit regarding the assumption Williamson actually employs. Second, the appeal to milliseconds doesn't quite fit the case as described. In this section, I'll be concerned with the first of these problems. In the next section, I'll discuss the second.

There are two ways of reading NAC: one utterly innocuous, the other highly controversial. Suppose the following is true:

(1) I'm not aware of any change.

The explanation of why (1) is true might simply be that no change has taken place. No such explanation is possible for:

(2) There has been a change of which I'm not aware.

If (2) is true, then a change has taken place, and it's a change of which I'm not aware. Likewise, if we consider:

NAC* In case α, there are no one-millisecond intervals of feeling less cold of which one is aware

there is the possibility that NAC* is true for the following reason: One is not aware of any one-millisecond intervals of feeling less cold because no such intervals have occurred. Lucy would find this possibility utterly unobjectionable. On the other hand, we might read NAC this way:

NAC** In case α, there are many one-millisecond changes of feeling less cold than before such that one is not aware of them due to their being indiscernible.

While a luminosity friend like Lucy need not reject NAC*, she will of course reject NAC**. Now, the premise on which Williamson's argument depends is not NAC* but NAC**. Without the assumption that indiscernible one-millisecond changes of feeling less cold than before actually occur, Williamson will not be able to justify

(P_R2) If K(cold) in α_i then B(cold) in α_{i+1}.[12]

But Williamson needs (P_R2) for his defense of (R). We must, therefore, interpret NAC as NAC**.

It is important to notice the difference between NAC* and NAC** because, when reading Williamson's anti-luminosity argument, one might easily interpret NAC as the innocuous NAC*. Not noting that the premise actually at work in Williamson's argument is NAC**, one might overlook that NAC** is a controversial metaphysical claim. As a result, one might ascribe more strength to the anti-luminosity argument than it really possesses.

Warming Up Slowly and Warming Up Quickly

Let us continue with the second problematic feature of NAC: the choice of milliseconds. For a fast-moving thing such as light, we use a short unit of measurement: the speed of light is 299,792,458 meters per *second*. For comparatively slow moving objects like cars, cyclists or pedestrians, we measure their speed in miles or kilometers per *hour*. Suppose a car crashes into a building, coming to an almost instant stop. Since its slowdown is so *fast*, we probably want to measure deceleration in milliseconds. In contrast, suppose a 100 car freight train loses power and slows down very *slowly*, taking about an hour to come to a complete stop. Since it decelerated so slowly, there is no need to measure its deceleration in milliseconds. Minutes will do just fine. Now, in Williamson's example, one warms up very *slowly*. Since one warms up only slowly, it's just not plausible to assume that in case α there are any one second, let alone any *one-millisecond* changes of feeling less cold than before. Indeed, since one's warming up is stretched out over several hours, it's not even plausible to assume that there are one-minute changes of feeling less cold than before.

[12] Why, if case α won't involve indiscernible one-millisecond changes of feeling less cold, will Williamson not be able to justify (P_R2)? Case α is a case in which one gradually warms up; it must be viewed as a long series of changes of feeling less cold than before. Now, if we think of an α_i - α_{i+1} interval as possibly encompassing a change of feeling less cold than before such that one is *aware* of this change, then we get the possibility of K(cold) in α_i and ~B(cold) in α_{i+1}. To block this possibility, Williamson must assume that the change of feeling less cold occurring during an α_i - α_{i+1} interval is such that one cannot be aware of it. Hence Williamson needs NAC**.

The assumption that one undergoes one-millisecond changes of feeling less cold is more plausible when we consider a case of a short duration in which one warms up very *quickly*. If I go to my mailbox on a cold winter day in Minnesota when it's -10°F, I get very cold even though I'm outside for just a couple of minutes. When I step back inside the well-heated house, warming up again also just takes a couple of minutes. Since once back in the house I warm up quickly, it might be that I feel less cold in such rapid increments that my awareness of the process lags behind. So while I'm aware of one one-second changes of feeling less cold, perhaps there are many one-millisecond changes of feeling less cold of which I am unaware.

The use of milliseconds is essential for Williamson. Why think that an episode of feeling less cold than before might not be noticeable to one? Answer: It won't be noticeable when it lasts for only one-millisecond. So let's focus on the plausibility of one-millisecond changes of feeling less cold than before.

Feeling Less Cold in One Millisecond

Williamson, I take it, considers NAC an uncontroversial premise, a mere stipulation that even a luminosity friend such as Lucy should accept. Let us examine whether this is so. Lucy would say that the issue is not empirical but a conceptual or metaphysical. Metaphysically, feeling less ϕ and being aware of feeling less ϕ cannot come apart, just like one's feeling ϕ and one's *being alive*, or one's feeling ϕ and one's *being conscious* cannot come apart. These things cannot come apart whether we consider minutes, seconds, or milliseconds. What matters is not the duration of the interval but the metaphysical nature of the episodes in question: episodes of feeling less ϕ than before not accompanied by any awareness. According to Lucy, episodes of that nature are metaphysically impossible. Choosing an extremely short interval doesn't change that. If an episode of feeling less ϕ without awareness is metaphysically impossible, it's still metaphysically impossible if it's supposed to last for only one millisecond. Lucy will deny, therefore, that the choice of milliseconds supplies her with a reason to consider NAC** plausible.

I must confess I find Lucy's position more plausible than Williamson's. When I put a pot of water on my stove, there might be one-millisecond increases in the water's

temperature. When I put the pot of boiling water on my porch when it's -20°F, there might be one-millisecond decreases in the water's temperature. But *temperature* increases and decreases are one thing; a person's *feeling* more or less hot, or more or less cold, than a moment before is another. So the possibility of one-millisecond temperature changes provides no support for the claim that one can feel less cold than before within the span of one millisecond.

Let's switch to other mental states. Are there indiscernible episodes of feeling less pain, indiscernible episodes of feeling less tired, or indiscernible episodes of feeling less happy? I would have to agree with Lucy's view. F*eeling* less pain, less tired, or less happy is, by its very nature, not the sort of thing that can take place without awareness of it. I see no reason to suppose that this is different for feeling less cold.

Suppose we grant the metaphysical possibility of feeling less ϕ without being aware of feeling less ϕ. Will we then have to attribute plausibility to the claim that it's possible for one to undergo *one-millisecond* episodes of feeling less cold? Here's a reason to think not. Consider the Minnesota Winter case. Having stepped back inside the well-heated house, I warm up very quickly. Since I warm up very quickly, it's perhaps plausible to say that each *minute* constitutes a change of feeling less cold than before. Perhaps we should even allow for *twenty second* changes of feeling less cold than before. But do I also undergo *millisecond* changes of feeling less cold? What reason is there to think I do? There is only one candidate: my *warming up rapidly*. But my warming up rapidly is not a compelling reason for thinking that I undergo millisecond changes of feeling less cold.

Let us compare two propositions:

(T) In the Minnesota Winter case, I undergo *twenty-second* changes of feeling less cold.

(M) In the Minnesota Winter case, I undergo *one-millisecond* changes of feeling less cold.

My warming up rapidly is an important feature of the Minnesota Winter case. I think it is a reason to accept (T). But I do not think it is also a reason to accept (M). In granting (T), we are already taking into account that the Minnesota Winter case is one in which one warms up rapidly. Why grant (M) as well? If we are to view (M) as plausible, we need to

be supplied with an *additional* reason. I don't find it easy to imagine what that reason might be.

What about *micro*seconds? Would Williamson want to say that there are also one-microsecond episodes of feeling less cold, or would he want to draw the line at milliseconds? If he does *not* hold that there are one-microsecond episodes of feeling less cold, then his position will be this:

> It's possible to undergo episodes of feeling less cold within the span of one one-thousandth of a second, but it it's not possible to undergo episodes of feeling less cold within the span of one one-millionth of a second.

This position strikes me as a bit arbitrary. Is there any principled reason, other than the limits of one's own awareness, for marking a threshold beyond which an experience like feeling less cold than before can no longer occur? If there is such a reason, what might it be? Suppose there is a good reason for saying that microseconds of feeling less cold don't exist. If so, why wouldn't that reason also be a good reason for saying that milliseconds of feeling less cold don't exist?

Taking on Lucy

Lucy holds the equivalence $F_L\phi \leftrightarrow A(F_L\phi)$. Therefore, she has no reason to accept NAC: the assumption that in case α, there are episodes of feeling less cold than before of which one is not aware. As a result, Williamson's argument is not dialectically effective against Lucy. Lucy might say that the argument begs the question against her.

We can also look at the dialectical situation this way. Williamson's aim is to show that no mental conditions are luminous. To this end, he employs a reductio-type argument resting on the premise that a mental condition such as feeling cold is luminous. But in order to derive the contradiction the reductio requires, he presupposes a specific kind of non-luminosity: namely the non-luminosity of feeling less cold than a moment before, where the moment in question is one millisecond. For two reasons, this assumption is problematic. First, it is unclear whether it is psychologically possible to feel less cold than before within the span of one millisecond. Second, Lucy will see no reason to grant that, whereas as feeling cold is luminous, feeling less cold than a moment before is not

luminous. Consequently, Lucy will not think that Williamson's argument supplies her with a reason to abandon luminosity.

For Williamson's argument to be effective against a luminosity friend like Lucy, he would have to establish some common ground between Lucy and himself. The appeal to milliseconds, I take it, is intended to find such common ground. So starting with the assumption that one-millisecond changes of feeling less cold than before are indiscernible, Williamson attempts to give luminosity friends a reason to abandon luminosity. However, since Lucy accepts the equivalence $F_L\phi \leftrightarrow A(F_L\phi)$, she has no reason to believe that one-millisecond changes of the kind in question are metaphysically possible. Her argument goes as follows: "If one-millisecond changes of feeling less cold are psychologically possible, they would have to be discernible. If, on the other hand, the human mind is such that one-millisecond changes are not discernible, then there cannot be one-millisecond changes of feeling less cold." The appeal to milliseconds does not, therefore, establish the common ground needed for supplying Lucy with a good reason to abandon the luminosity of feeling cold.

A Dilemma

Consider again the following premise which Williamson needs for his defense of (R):

(P_R2) If K(cold) in α_i then B(cold) in α_{i+1}.

Why believe that, in case α, there won't be two times such that K(cold) in α_i and ~B(cold) in α_{i+1}? Williamson's answer, I take it, is that α_i and α_{i+1} are just one millisecond apart, and within the span of one-millisecond one is not going to be aware of feeling less cold than a moment before. Suppose we allow for the $\alpha_i - \alpha_{i+1}$ interval to be long enough to notice that one feels less cold than before. Then surely there will be a pair of times such that K(cold) in α_i and ~B(cold) in α_{i+1}: in α_i one still feels cold and knows that one does, but in α_{i+1} one's feeling cold has passed and so one no longer believes one feels cold. Indeed, given that in case α one moves from feeling cold to eventually feeling hot, and given that one constantly monitors how one feels, there will have to be such a pair. So if we permit the α_i - α_{i+1} interval to be long enough for there to be awareness of feeling less cold, then (P_R2) is false and (R) cannot be derived. That's why letting the $\alpha_i - \alpha_{i+1}$ interval be no longer than one millisecond is essential to Williamson's argument.

However, as I have argued above, Williamson assumes that in case α, one warms up very *slowly*. Suppose we disagree with Lucy and allow for the metaphysical possibility of indiscernible episodes of feeling less cold than before. Rejecting Lucy's view is not enough to render plausible the assumption that in case α there are one-millisecond intervals of warming up. For in case α as described by Williamson, one warms up *slowly*. That's an excellent reason to think that in case α there are no one-millisecond warm-up episodes. What might be psychologically realistic? Considering that the warming-up process takes several hours, perhaps it's plausible to assume that one feels less cold every minute. But if we think of the $\alpha_i - \alpha_{i+1}$ interval as lasting one minute, then surely there is no reason to think that one's coming to feeling less cold during this interval is indiscernible. So if we think of case α in terms of one-minute increments of warming up, nothing blocks the possibility of K(cold) in α_i and ~B(cold) in α_{i+1}. That's the first horn of the dilemma.

Here's the second horn. To block the possibility of K(cold) in α_i and ~B(cold) in α_{i+1}, Williamson needs to make the $\alpha_i - \alpha_{i+1}$ interval as short as it realistically might be. Now, to make it plausible that one undergoes very short episodes of feeling less cold than before, we need a case in which one warms up rapidly. But even in the Minnesota Winter case, one still does not warm up quickly enough for it to be plausible to think that the case is going to involve one-millisecond episodes of feeling less cold than before. Let's be charitable and allow for one-second episodes of feeling less cold than before. But a one-second episode of warming up still seem long enough to go from K(cold) to ~B(cold).

Let's consider what might be the fastest possible case of warming up. Having been outside during a brutally cold Minnesota winter night, I come back home, immediately undress and get into my bathtub, which is filled with pleasantly warm water. Let's suppose upon immersing myself in the warm water, I do indeed feel less cold from one millisecond to the next.[13] Compared with the original case Williamson considers in which one warms up slowly, we are now looking at the other end of the spectrum. But why should we assume that, while luxuriating in my bathtub, the one-millisecond warm-up

[13] Getting into the tub will take longer than one millisecond. So let's suppose the series of one-millisecond changes of warming up begins once I'm fully immersed in the warm water.

episodes I enjoy are not accompanied by awareness? If they are, we get the pair K(cold) in α_i and ~B(cold) in α_{i+1} even though the $\alpha_i - \alpha_{i+1}$ interval lasts only one millisecond.

Suppose you are in a pitch dark room. Someone turns on the light. It seems plausible to me that here we have K(dark) at t_i and ~B(dark) at t_{i+1} where t_i and t_{i+1} are only one millisecond apart. Or suppose it's pleasantly quiet now but then a screeching smoke alarm goes off. Again, I don't see why we shouldn't think that you can pass within one millisecond from K(quiet) to ~B(quiet). Likewise, if we assume maximum warm-up speed and make the contrast between feeling cold and feeling less cold maximally extreme, why shouldn't we assume that a one-millisecond change of feeling less cold than before is accompanied by awareness?

So what's the dilemma? The two horns of the dilemma are (i) cases in which one warms up very slowly and gradually and (b) cases in which one warms up very quickly and suddenly. If the warming-up process is gradual, there's no reason to believe that the individual increments of warming up are very short. If they are not very short, it's not plausible to assume they are indiscernible. If, on the other hand, the warm-up speed is fast enough to render one-millisecond changes plausible, there's then again no reason to think that the warming-up increments are indiscernible. It is now the very speed of the warm-up process that gives us a reason to think that even one-millisecond changes might be discernible. So no matter whether we consider a slow and gradual warm-up cases or a fast and sudden warm-up case, it would seem we get the possibility of K(cold) in α_i and ~B(cold) in α_{i+1}. As a result, even if we don't agree with Lucy's metaphysics, we may wonder why we should accept Williamson's reasoning in support of (R).

Being Appeared to ϕ-ly

Let's be done with mental states such as feeling less ϕ. Instead, let's consider mental states such as *being appeared* to ϕ-ly. Williamson holds that the following is metaphysically possible: It appears to one that one is in the state of being appeared to ϕ-ly while one is actually in the state of being appeared to ϕ*-ly. Williamson describes the following case:

> Suppose that Jones has a clear and distinct experience E as of seeing 29 stars. On the basis of having E, Jones forms the belief that he is having an experience as of

> seeing 29 stars . . . However, for any natural number *n* between 20 and 40, when Jones has an experience as of seeing *n* stars, he usually forms a belief that he is having an experience as of seeing 29 stars. In most cases, this belief is false. The underlying psychological mechanism is the same for all those values of *n*. He makes no attempt to count but simply estimates the number from his general impression; forgotten events in his childhood caused a strong bias in favor of the number 29.[14]

So according to Williamson, it's possible for Jones to believe himself to have an experience as of 29 stars when in fact he as an experience of, say, 30 stars. But is this really possible?

Let's distinguish between type-one deception and type-two deception. Type-one deceptions occur when one's perceptual faculties represent one's *physical environment* other than it really is. Type-two deceptions occur when one's introspective faculties represent one's own *mental states* other than they really are. Modelled after Williamson's example, here is an illustration of a type-one deception: It seems to me that I'm seeing 29 stars, but I'm actually seeing 30 stars. The number of stars in my visual field is one thing; my experience of them is another. There is no necessary connection between them. Therefore, the possibility of a type-one deception like that is entirely unproblematic. When a type-two deception occurs, we have again an experience that represents an external reality other than it really is. But in this case, the external reality is one's own mind, and the misleading experience is of the introspective kind. Using Williamson's example involving the perception of stars in the sky, we might want to think of the two types of deception as follows:

Type-One Deception

Reality:	There are 30 stars in the sky.
Visual Appearance:	It visually seems to me that there are 29 stars in the sky.

Type-Two Deception

Reality:	It visually seems to me that there are 30 stars in the sky.
Introspective Appearance:	It introspectively seems to me that it visually seems to me that there are 29 stars in the sky.

[14] Williamson 2005, p. 471.

According to Williamson, introspective deception about how the external world appears to one in one's perceptual experiences is just as possible as perceptual deception about what the external world outside of one's mind is really like.

But why should we assume that type-two deceptions are possible? Well, Williamson has described an example that enjoys at least some initial plausibility. I will argue, however, that neither Williamson's example nor any other example of that kind is likely to give us a compelling reason for the metaphysical possibility of type-two deceptions.

Consider an assertion of the form "It's possible for one to believe one is appeared to ϕ-ly when in fact one is appeared to ϕ*-ly." To acquire a reason to believe that what this assertions alleges to be possible is actually possible, we need to be given *details* that help us understand exactly *how* such a deception about one's own mental states can come about. Williamson supplies us with details meant to accomplish this task. They are as follows (switching from Jones to myself as the subject):

> **The Details (D)**
>
> Forgotten events in my childhood have caused in me a strong bias in favor of the number twenty nine. This bias kicks in whenever I perceive *n* objects, where *n* is a natural number between 20 and 40.

But (D) does not force us to interpret Williamson's example as involving type-two deception. Note that, when it comes to interpreting what's going on in Williamson's example, we have two options. We could say that, given (D), the example is plausibly viewed as one involving type-two deception. However, (D) allows equally well for interpreting the example as involving type-one deception. The example will give us a reason to believe in the metaphysical possibility of type-two deception only if (D) is incompatible with type-one deception, thus forcing us to interpret the example as an illustration of type-two deception.

I can detect no such incompatibility. Given (D), we might say that even though there are 30 stars visible in the sky, it seems to me that there are 29 stars in the sky, not because I have counted but because of my bias in favor of the number 29. Of course, the example *could* also be interpreted as involving a type-two deception. But to a luminosity friend such as Lucy, such an interpretation is bizarre, conflicting with the following intuition: necessarily, if it introspectively appears to one that one is appeared to ϕ-ly, then one *is*

appeared to ϕ-ly. Thus, if it seems to me (introspectively) that it seems to me (visually) that there are 29 stars in the sky, then it (visually) seems to me that there 29 stars in the sky. If one shares this intuition with Lucy, then one will not easily agree that Williamson's example represents a genuine metaphysical possibility. One will agree that the example is metaphysically possible only if one is given details which mandate interpreting the example as involving a type-two deception. However, (D) does not require us to adopt such an interpretation. I do not think, therefore, that Williamson's star example gives us a reason to doubt the luminosity of being appeared to ϕ-ly.

Luminosity and Evidential Defeat

In a symposium on Williamson's *Knowledge and its Limits*, Earl Conee has proposed an argument against the luminosity of mental states that is based on the possibility of evidential defeat.[15] Conee assumes the following premise:

D1 If someone has evidence against *p* that is strong enough to defeat whatever grounds the person has for believing that S is true, then the person is insufficiently justified to know that *p* is true.[16]

Suppose I believe I feel cold. What justifies this belief? There are no uncontroversial answers to this question. For the sake of discussion, let's settle on the following answer, which will work well for discussing Conee's argument: What justifies my belief is that it (introspectively) appears to me that I feel cold: Ap(cold) justifies B(cold). According to Conee, Ap(cold) is vulnerable to defeating evidence, even if we "reject a decidedly dubious distinction between how things seem to be experienced by someone, and how they really are experienced."[17] It would appear Conee is sympathetic to Lucy's view, which includes as well the following equivalence: One's feeling Φ always goes together with it appearing to one that one feels Φ. One cannot occur without the other. So: FΦ ↔ Ap(FΦ). Conee's argument appears to be that even granting this equation, it is still possible for one to feel FΦ while having evidence for believing one does not feel Φ. Conee says:

[15] Reference.
[16] Conee 2005, p. 448.
[17] Ibid.

> For instance, a defeating argument can appeal abstractly to experts who are alleged to have discovered a very subtle difference here. Or it can be alleged that introspection gives us access to our phenomenal qualities via modes of representation of those qualities, and it has been discovered how to stimulate a brain so as to duplicate such modes in the absence of the quality.
>
> In fact, no explanation of the possibility is required. Sheer testimony can be strong enough. For instance, one who is known by Smith to be an expert can quite credibly assert to Smith that merely apparent phenomenal qualities are possible, without explaining how this can be . . .
>
> Suppose Smith is given some . . . reason to doubt that she is experiencing a certain chilly feeling, and she has nothing available to her to refute it. The reason then defeats the evidence provided by her experience of the chilly feeling. With her evidence defeated, we can infer by D1 that she does not know herself to experience that quality.[18]

Of course, if a mental condition C is luminous, it need not be the case that whenever C obtains one knows that C obtains. Rather, if C is indeed luminous, it would have to be the case that, whenever C obtains one is in a position to know that C obtains. Hence Conee continues:

> Is [Smith] still in a position to know that she is in the condition [of experiencing a certain chilly feeling]? We began with Tim's characterization of the things that one is in a position to know as facts that are open to one's view, unhidden, and without any obstacle to one's knowing them. We now have reason to separate the first two features from the third. The fact of Smith's experiencing the feeling remains open to her view and unhidden. But her justification for believing herself to experience the feeling has been defeated. This is an obstacle to her knowing the fact of her experiencing the feeling and, absent new evidence, it is an obstacle that she cannot surmount.[19]

Williamson responds to Conee's reasoning noting that he has "considerable sympathy for Conee's anti-luminosity argument." Williamson anticipates a friend of luminosity to

[18] Conee 2005, p. 448f.
[19] Ibid.

reply that, "despite the defeaters, the agent is still in a position to know, and perhaps would know if she ignored the defeaters. However, Williamson adds: "I do not endorse such a reply."[20]

Well, let's have a look at what we might call the *indefeasibility view*. Suppose Ingrid, a friend of Lucy's, endorses that view. What might Ingrid say in response to Conee's argument? She might reason as follows:

> **Ingrid's Indefeasibility View**
>
> Suppose you have a splitting headache. Can an "expert" really defeat your evidence for believing you have a headache, thus preventing you from knowing you have a headache? Not really. You are aware of your headache. Your headache hurts. Such evidence defeats all defeaters. The "expert" might tell you a story about the latest research in cognitive psychology. He might show you a faculty ID from one of the most prestigious universities in the country. But now matter what he says, he cannot succeed in defeating your awareness of your headache as evidence for believing that you have a headache. Since you are aware of your headache, you have an excellent reason to believe that, on this occasion, the seemingly renowned "expert" is mistaken. You remain, in spite of what the expert says, in a position to know that you have a headache. Analogous reasoning applies to believing that you feel cold. If you feel cold, then you are aware of feeling cold. Such awareness defeats all defeaters. Whenever you have such evidence, you are in a position to know that you feel cold, no matter what an "expert" is going to tell you about your feeling cold.

Ingrid does not recommend that you *ignore* the expert's testimony. There's no need to ignore it since you have a defeater for it: your awareness of your mental state. What Ingrid recommends is to conclude, on the basis of your awareness, that the person who appears to be an expert got it wrong.

Is Ingrid obviously mistaken? Perhaps she's mistaken, but not obviously so. As for myself, I would have to say I find her view more plausible than the view Conee and Williamson endorse. Expertise has its limits. When an expert attempts to convince me that while I'm aware of feeling cold I am or might be mistaken in believing I feel cold,

[20] Williamson 2005, p. 473.

the limits of expertise have been reached. I'm still in a position to know I feel cold. Or so it seems to me. I do not think, therefore, that Conee's argument succeeds in undermining the luminosity of feeling cold.

Luminosity and Experiential Foundationalism

In addition to the truth of the matter, what is at stake in the debate over the luminosity of mental states? Consider experiential foundationalism, a theory placed well within the internalist framework that Williamson's *Knowledge and its Limits* systematically and thoroughly rejects.[21] We might wonder whether experiential foundationalism somehow depends on the luminosity of experiences.

Suppose I have a perceptual experience as of *p*: E*p*. According to experiential foundationalism, E*p* can, without the aid of any further beliefs, justify me in believing that *p*. From an internalist point of view, E*p* qualifies as a justifier because, if one has an experience as of *p*, one can tell that one does. That's why such an experience gives one a reason to believe that *p*. If experiences are luminous, then whenever E*p* occurs, one is in a position to know that E*p* occurs. Suppose, then, experiences are not luminous. If so, there are occasions when E*p* occurs, but one is not in a position to know that E*p* occurs. If there were such occasions, would that be a problem for experiential foundationalism?

Internalist foundationalism has been accused of breeding skepticism. The basic charge is that, if justification is to meet internalist accessibility constraints, justification becomes too difficult to come by.[22] Experiential foundationalism is not vulnerable to this objection. The view asserts that an ordinary experience as of *p* can justify a person in believing that *p*. Neither beliefs about one's experience, nor metabeliefs about one's belief that *p*, nor any appeal to any epistemic principles are required. The experience as of *p* by itself, if undefeated, is sufficient for justifying a person in believing that *p*. Moreover, the view qualifies as internalist because experiences are mental states, and as such differ from external reality in being . . . well, *luminous*. So if Williamson is right about the non-luminosity of mental states, then we might think that experiential

[21] See Feldman 2003, pp. 70-78 and 145ff, Huemer 2000, Pryor 2000, and Steup 2005.
[22] See Goldman 1999.

foundationalism no longer qualifies as an internalist theory. This, however, doesn't follow.

Suppose Williamson is right: mental states are not luminous. If they are not luminous, then they are *not always* recognizable. But they will still be recognizable *a lot of the time*. Experiential foundationalism can easily accommodate this result. Let's distinguish between restricted and unrestricted experiential foundationalism:

> **Unrestricted Experiential Foundationalism (UEF)**
>
> Whenever one has an experience as of *p*, one has internalist justification for believing that *p*.
>
> **Restricted Experiential Foundationalism (REF)**
>
> Whenever one has a discernible experience as of *p*, one has internalist justification for believing that *p*.

According to UEF, all experiences are justifiers; according to REF, only those that are discernible, that is, only those one is in a position to know one has them. UEF requires luminosity, for internalists would not want to view an indiscernible experience as of *p* as something that can justify a subject in believing *p*. REF, on the other hand, does not. By restricting justifying experiences to discernible experiences, REF accommodates the existence of indiscernible experiences. The latter, experiential foundationalists would say, are not a source of justification precisely because they are indiscernible.

So if we assume, hypothetically, that mental states are not luminous, then experiential foundationalists can retreat from UEF to REF. This wouldn't be much of a loss. After all, Williamson's attack on luminosity does not suggest that a huge percentage of our experiences are indiscernible. Attacking luminosity is not the same as making a case for universal skepticism about knowledge of one's own experiences. After all, indiscernible experiences occur only, if they occur at all, in the close vicinity of the transition point when one gradually changes from feeling ϕ to not feeling ϕ. Moreover, indiscernible experiences do not prompt beliefs whose justification needs to be accounted for. Suppose it appears to me that I feel ϕ, but actually I feel ϕ^* because my experience changes in a subtle and unnoticeable way from feeling ϕ to feeling ϕ^*. So my feeling ϕ^* is indiscernible to me. Well, if so, I will not believe I feel ϕ^*. Rather, I will still believe I feel ϕ. This belief, though mistaken, is justified. What justifies it is that it *appears to me*

that I feel ϕ, which is a discernible experience. Of course, this experience is misleading. It justifies me in believing something false. But the fallibility of experiential justification is just one more feature of experiential foundationalism, to be viewed as one of the theory's virtues. It's difficult to see, then, why we should think that abandoning indiscernible experiences as a source of justification should be a problem for experiential foundationalism.

However, the need for the retreat from UEF to REF arises only if it is true that mental states are not luminous. Of the three anti-luminosity arguments discussed here, each turned out to be problematic. I do not think, therefore, that they supply experiential foundationalist with any compelling reason to retreat to the restricted position.[23]

[23] I wish to thank Mylan Engel, Allen Habib, Omar Mirza, Duncan Pritchard, Casey Swank, and Kenneth Williford for helpful discussion and comments.

References

1. Brueckner, Anthony and Fiocco, Oreste. 2002. “Williamson’s Anti-Luminosity Argument.” *Philosophical Studies* 110, pp. 285-293.
2. Feldman, Richard. 2003. *Epistemology*. Upper Saddle River: Prentice Hall.
3. Goldman, Alvin. 1999. "Internalism Exposed." *The Journal of Philosophy* 96, pp. 271-293.
4. Huemer, Michael. 2000. *Skepticism and the Veil of Perception*. New York: Rowman and Littlefield.
5. Neta, Ram and Rohrbaugh, Guy. 2004. “Luminosity and the Safety of Knowledge.” *Pacific Philosophical Quarterly* 85, pp. 396-406.
6. Pryor, James. 2000. "The Skeptic and the Dogmatist." *Noûs* 34, pp. 517-49.
7. Steup, Matthias. 2004. "Internalist Reliabilism." *Philosophical Issues* 14, pp. 401-425.
8. Ramachandran, Murali. 2005. “Anti-Luminosity.” Unpublished manuscript available at http://www.sussex.ac.uk/Users/muralir/pubs/.
9. Williamson, Timothy. 2000. *Knowledge and its Limits*. Oxford: Oxford University Press.
10. Williamson, Timothy. 2005. “Replies to Commentators.” *Philosophy and Phenomenological Research* 70, pp. 468-491.