Philosophy and Phenomenological Research
10.1111/phpr.12067

PHPR | Journal Name | Manuscript No. | 1 2 0 6 7 | B
Dispatch: 7.8.13 | Journal: PHPR | CE: Mahalakshmi
Author Received: | No. of pages: 27 | PE: Vinoth

Are We Luminous?

AMIA SRINIVASAN
All Souls College, University of Oxford

Since its appearance over a decade ago, Timothy Williamson's anti-luminosity argument has come under sustained attack. Defenders of the luminous overwhelmingly object to the argument's use of a certain margin-for-error premise. Williamson himself claims that the premise follows easily from a safety condition on knowledge together with his description of the thought experiment. But luminists argue that this is not so: the margin-for-error premise either requires an implausible interpretation of the safety requirement on knowledge, or it requires other equally implausible (and soritical) assumptions. In this paper I bolster the margin-for-error premise against these attacks by recasting Williamson's own two-part defence, the first intended to work on the assumption that there is no constitutive connection between the phenomenal and the doxastic, and the second intended to work without this assumption. *Pace* various luminists, I argue that the appeals to safety needed for Williamson's two-part defence (the first in terms of outright belief, the second in terms of degrees of confidence) are plausible. I also argue that all that is needed to generate the margin-for-error premise from these safety conditions is an *empirical* assumption about the kinds of creatures we are: that is, creatures whose beliefs are structured by certain dispositions. By recasting the anti-luminosity argument in this way, we can understand what is really at stake in the debate about luminosity: that is, whether *we* are luminous.

1. Introduction

In *Knowledge and Its Limits*, Timothy Williamson argues that there are no non-trivial luminous conditions, where a condition is luminous just in case whenever one is in it, one is in a position to know one is in it (2000, chp. 4). If Williamson is right, then the common picture of the phenomenal realm as one of privileged access turns out to be a Cartesian orthodoxy from which philosophy must be cleansed. It also follows that rationality, evidence, normative obligations, and sameness of meaning – phenomena associated, for many, with privileged first-person access – are themselves non-luminous.

Given its potential to destabilise, it is little wonder that the anti-luminosity argument has come under fire since its appearance. Luminists typically attack Williamson's use of a certain *margin-for-error premise*. Williamson himself claims that the premise follows easily from a safety condition on knowledge together with his description of the thought experiment. But luminists argue that this is not so: the margin-for-error premise either requires an implausible interpretation of the safety requirement on knowledge, or it requires a plausible interpretation of the safety requirement together with other implausible, often soritical, assumptions (Weatherson 2004; Blackson 2007; Wong 2008; Berker 2008; Ramachandran 2009; Vogel 2010; Cohen 2010; Zardini forthcoming). Either way, the margin-for-error premise, and thus the anti-luminosity argument, is in trouble. Luminists counsel that we dismiss Williamson's argument and cleave to the luminous.

I shall argue that Williamson's controversial margin-for-error premise, *pace* the luminists, can be derived from a plausible safety condition on knowledge[1] together with a plausible empirical hypothesis about the kind of creatures we are – creatures, namely, whose beliefs are structured by certain kinds of dispositions. Indeed, I shall make this argument twice over. This is because some luminists have been keen to argue that the margin-for-error premise is particularly problematic on any view that maintains a constitutive connection between the phenomenal and the doxastic. So I will first argue for the margin-for-error premise from safety and empirical considerations on the assumption that no such constitutive connection obtains. I will then make a more refined argument that dispenses with that assumption: that is, an argument that applies to all (non-trivial) phenomenal conditions regardless of any constitutive connection that might obtain between them and beliefs about whether they obtain. In so doing, I hope to show the anti-luminosity argument to be robust against some of the most common criticisms. I also hope to show that a large part of what is at stake in the debate about anti-luminosity is a certain vision of what kind of creatures we are, empirically speaking.

[1] I won't be addressing those luminists who reject safety wholesale (Brueckner & Fiocco 2002; Neta & Rohrbaugh 2004; Comesaña 2005; Conee ~~200~~), only those who accept a safety condition on knowledge while objecting to the particular versions of it (putatively) needed for Williamson's argument.

2. The Anti-Luminosity Argument

Williamson aims to establish that, for almost any condition,[2] it is possible for a normal human to be in that condition and fail to be in a position to know that she is in it.[3] He attempts to do this by producing a counterexample to the putative luminosity of the condition of *feeling cold* – a condition that, to many at least, seems paradigmatically luminous. Since analogous thought experiments can be produced for any other putatively luminous (non-trivial) condition, the anti-luminosity argument should generalise to all (non-trivial) conditions. Here is the thought experiment:[4]

> **Cold Morning**. S wakes up at dawn feeling freezing, very slowly warms up, and feels hot by noon. Throughout the morning S is concentrating sufficiently hard on the question of whether she feels cold, such that if she is in a position to know that she feels cold then she does indeed know. S's powers of discrimination are limited, and the change from S's feeling cold to hot is so gradual that S "is not aware of any change in them over one millisecond" (Williamson 2000, p. 97). S's confidence that she feels cold gradually diminishes, such that by noon she firmly believes that she no longer feels cold.

Let t_0, t_1, $t_2 \ldots t_n$ be a series of times at one-millisecond intervals from dawn to noon. Let α_i be the case[5] at time t_i. Let C be the condition *that S feels cold*, and K(C) the condition *that S knows that C obtains*.

Now, let us assume that C is a luminous condition for S. That is, whenever S is in C, she is in a position to know she is in C. By the description of Cold Morning, whenever S is in a position to know that C obtains, she does in fact know that C obtains. Thus we have:

> **(LUM)** If C obtains in α_i then K(C) obtains in α_i

---

[2] Williamson notes that some conditions might not be susceptible to his anti-luminosity argument; such conditions might be *trivially* luminous. For example, conditions that never obtain are vacuously luminous, and conditions that always obtain might be luminous (when presented under certain guises). Williamson's argument also does not work against eternal conditions, which always obtain if they ever obtain, though he gestures at an argument that shows that even such conditions are not plausibly luminous (2000, p. 108). Williamson's central point is not that there could be no luminous conditions, but rather that luminous conditions, if they exist, are "curiosities" (ibid, p. 109). Those conditions that we think of as paradigmatically luminous – e.g. being in pain, feeling cold, having a desire to f, seeming/appearing that p – are, if Williamson is correct, non-luminous.

[3] That is, it is possible for her to introspect as assiduously as possible without thereby coming to know that she is in the condition.

[4] This is my description of the thought experiment, though it is similar to Williamson's original.

[5] A 'case' is a centred possible world – that is, a possible world with a designated subject and time.

Williamson then introduces the following *margin-for-error* principle, which he claims falls out of a simple safety condition on knowledge together with the description of the Cold Morning:

(MAR)[6] If K(C) obtains in α_i then C obtains in α_{i+1}

By the description of Cold Morning, at dawn S feels cold, and at noon she no longer feels cold. So we have:

(BEG) C obtains in α_0

(END) C does not obtain in α_n

(LUM), (MAR), (BEG) and (END) are together incompatible. By (LUM), if C obtains in α_0, then S knows that C obtains in α_0. By (MAR), if S knows that C obtains in α_0, then C obtains in α_1. By (BEG), C does obtain in α_0; therefore, C obtains in α_1. Similarly, we can establish that C also obtains in α_2, α_3, $\alpha_4 \ldots \alpha_n$. So C obtains in α_n. But according to (END) C doesn't obtain in α_n. Thus we arrive at a contradiction.[7]

3. The Problem with (MAR)

Since (BEG) and (END) simply follow from the description of COLD MORNING, it seems we must either give up (MAR) or (LUM). Williamson counsels that we hold onto (MAR) and reject (LUM) – which is to say, abandon luminosity. Luminists, however, think that the lesson to be learned is that we should be suspicious of (MAR). But on what grounds?

Wong (2008) argues that (MAR) is derivable from the following two premises: (1) If in α_i one knows that one feels cold, then in α_i one feels cold, and (2) If in α_i one feels cold, then in α_{i+1} one feels cold. (1) follows uncontroversially from the factivity of knowledge, but (2) is of course soritical. Thus Wong concludes that (MAR) is itself soritical and should be rejected in favour of (LUM). But Wong is wrong to think that this obviously soritical argument is the only way to defend (MAR). In particular, Wong's defence of (MAR) does not appeal to a safety condition on knowledge, which Williamson clearly intends to be part of the justification for (MAR).

[6] There is no standard presentation of Williamson's argument in the literature, so (MAR), or its analogues in terms of possible worlds and times (rather than in the terminology of world-bound cases), is variously labelled (I_i) (Williamson (2000), Weatherson 2004; Blackson 2007; Ramachandran 2009); (R) (Vogel 2010); (C) (Wong 2008); (1) (Cohen 2010); and (KMAR) (Zardini forthcoming). I borrow (MAR) from Berker (2008).

[7] My presentation of Williamson's argument owes much to Berker's (2008).

Weatherson (2004) and Vogel (2010) offer safety conditions on knowledge from which we can directly derive (MAR). According to both Weatherson's 'content safety' condition and Vogel's 'strong reliability' condition, one knows that a condition R obtains only if R obtains in all very similar cases. As Weatherson and Vogel argue, this version of the safety condition directly secures (MAR), but is itself implausible. Intuitively, for S's belief that R obtains to be sufficiently safe for knowledge, it must be that there are no very similar cases in which S has an untrue belief that R obtains.[8] But a more plausible understanding of the safety condition[9] –one that only mandates no nearby untrue belief – is insufficient to directly derive (MAR). For (MAR) states that it is a necessary condition on S's knowing that C obtains in α_i that C also obtain in α_{i+1}. But S's belief that she feels cold in α_i could satisfy the 'no nearby ~~false~~ belief' condition so long as in all of the sufficiently similar not-C cases, S didn't *believe* that C obtained (cf. Berker (2008), p. 6). Luminists are right, then, to point out that a brute appeal to safety won't alone secure (MAR). And they are also right to point out that Williamson's anti-luminosity argument cannot run with safety alone; (MAR) is needed as a bridge principle from one moment to the next in order to deliver the *reductio* that S feels cold at noon. Thus the question is: how do we motivate (MAR)?

Williamson can be read as offering not one but two answers to that question, each relying on different specifications of the safety requirement on knowledge. The first defence, which employs a safety condition in terms of outright belief, which I call (BELIEF-SAFETY), is intended to work on the assumption that there is no constitutive connection between feeling cold and believing oneself to feel cold. The second defence, which employs a safety condition in terms of degrees of confidence, which I call (CONFIDENCE-SAFETY), is intended to allow for the possibility of such a connection. That is, the second defence is meant to secure (MAR), and thus the anti-luminosity argument, regardless of any possible constitutive connection between the phenomenal and doxastic realms. My plan is to elaborate and bolster each of Williamson's defences of (MAR) in turn. In the first instance, I will argue that all we need to add to (BELIEF-SAFETY) to generate (MAR) is a plausible empirical supposition about what kind of creatures we are. In the second, I will counter accusations that Williamson's (CONFIDENCE-SAFETY) is an implausible condition on knowledge, and

[8] I am assuming here and throughout my discussion of safety that for two cases of belief to be 'sufficiently similar' they must involve sufficiently similar *methods* of belief-formation. Some specifications of safety factorize 'sufficiently similar' into a modal component ('nearby') and a methods or basing component.

[9] Like Weatherson's 'belief safety' condition and Vogel's 'moderate reliability' condition.

argue that, again, (MAR) can be derived from it together with a plausible assumption about our empirical character.

4. Defending (MAR): Non-Constitutive Accounts

Williamson's preliminary defence of (MAR) is intended to work on the assumption that there is no constitutive connection between the phenomenal and the doxastic – specifically, that one's feeling cold isn't constitutively tied to believing one feels cold. This defence invokes a safety condition in terms of outright belief, which we can approximate as follows:

> **(BELIEF-SAFETY)** In case α S knows that a condition R obtains only if, in all sufficiently similar cases in which S believes that R obtains, it is true that R obtains.

Roughly, (BELIEF-SAFETY) says that knowledge requires not just true belief, but the absence of nearby untrue belief.[10] (BELIEF-SAFETY) is intuitively plausible. Imagine I look through the window and form the true belief that it's raining outside. Unbeknownst to me, a prankster has placed a screen outside my window that projects an image of rain. Clearly I don't know it's raining outside, though it is. This is because in a nearby world – the world in which it has just stopped raining – I have the false belief that it's raining.

Despite the plausibility of (BELIEF-SAFETY), Vogel (2010) argues that we should reject it in favour of what he calls 'moderate reliability' (p. 549), which amounts to this:

> **(VOGEL-SAFETY)** In case α S knows that a condition R obtains only if, in all sufficiently similar cases in which S believes that R obtains, it is not false that R obtains.

[10] What counts as a 'sufficiently similar case' (or a 'nearby world' or 'similar method') in definitions of safety is a vexed issue, analogous to what is known as the 'generality problem' for reliabilism (Conee and Feldman 1998). Williamson (2000) argues that the upshot of this problem is that we can offer no reductive analysis of reliability, and that our judgments about similarity of cases must be informed by our intuitions about what constitutes an instance of knowledge or ignorance. This means that any claim to knowledge or ignorance is subject to dismissal via an alternative judgment about similarity of cases. It is worthwhile noting that this is equally true of the knowledge/ignorance claims involved in Williamson's anti-luminosity argument. That is, the luminist could simply argue that the possible not-cold case in which S believed she was cold is too dissimilar to undermine the safety of S's belief in α_i. In particular, the luminist could simply argue that S's possible false belief has a different basis from S's actual true belief (in the first case, S believes *on the basis of feeling cold* that she feels cold; in the latter case she believes on some other basis). These are both easy (if unconvincing) ways of defending luminosity against Williamson's putative counterexample. Cf. Weatherson (2004, 4).

The difference between (BELIEF-SAFETY) and (VOGEL-SAFETY) comes into play when (if ever) it is neither true nor false that R obtains.[11] Suppose that as S moves through COLD MORNING, it is first true that S feels cold, then neither true nor false that S feels cold, and finally false that S feels cold. Now imagine that S is in the final instance of feeling cold, t_c. Can S know that she feels cold at t_c? (BELIEF-SAFETY) says no, while (VOGEL-SAFETY) says yes. Which safety condition gives us the correct treatment of cases such as these? Vogel offers the following thought experiment to motivate his version of safety:

> **Umpire**. Imagine that there is an umpire who is invariably correct about every clear case of balls and strikes. That is, whenever television replay can discern one way or the other, the umpire is right, even on extremely close pitches. Every once in a while, the umpire calls as a ball a pitch that seems "too close to call" even on replay (549).

Vogel's intuition is that the umpire is able to know about every clear case of balls and strikes, despite the fact in some of those cases he has nearby untrue beliefs. This result is compatible with (VOGEL-SAFETY) but not (BELIEF-SAFETY). My own intuition is that the umpire doesn't know in all the clear cases, since whether or not he has any nearby false beliefs, he sometimes has nearby *mistaken* beliefs – that is, he sometimes believes pitches to be balls when they aren't.[12] To my mind at least, the proximity of nearby beliefs that simply aren't true seems sufficient to destroy knowledge. No doubt these issues are murky. In a footnote, Vogel offers another thought experiment:

> **Color Chip**. You see a number of color chips. Some are perfectly red, while the others are borderline red. The chips are placed in an urn, and one chip is chosen at random. Before you see the outcome, you believe that the selected chip will be red simpliciter, and it happens to be perfectly red (p. 556, n. 15).

Here, as Vogel himself agrees, it seems that you can't know that the randomly chosen chip is red; this supports (BELIEF-SAFETY) over

---

[11] See Hawthorne (2005), Williamson (2005) and Zardini (forthcoming) for a discussion of the anti-luminosity argument and its relation to the phenomena of indeterminacy and vagueness.

[12] The defender of (VOGEL-SAFETY) might protest, as an anonymous referee did, that it's simply unfair to call such beliefs mistaken, since in these cases the proposition *that pitch is a ball* is neither true nor false. Might such cases not be a matter of 'spoils to the victor'? Perhaps we bottom out at intuitions here. All I can say is that I have ~~some~~ sympathy for the pitcher who walks a batsman on a pitch that is not in fact a ball~~,~~ and ~~then~~ laments that the umpire was mistaken. The batsman would be wrong to think the pitcher blameworthy for his call, but not wrong, I feel, to think his belief mistaken.

(VOGEL-SAFETY). It's also worth noting that (VOGEL-SAFETY) might have counterintuitive implications in cases of failed demonstratives.[13] Consider:

> **Phantom tollbooth**. You see a tollbooth, point to it and say: "That's a tollbooth". However, you're on a drug that makes you hallucinate all sorts of things; in a very nearby world, the drug causes you to hallucinate a tollbooth.

Intuitively, you don't know that what you're pointing to is indeed a tollbooth. But in the nearby possible world in which you hallucinate a tollbooth, the demonstrative 'that' fails to refer to anything at all, since there is no tollbooth. In such a possible hallucinatory case it seems intuitive to say that your tollbooth-belief is neither true nor false.[14] This result is delivered by (BELIEF-SAFETY) but not (VOGEL-SAFETY). Finally, take a modalised version of Kripke's contingent liar cases:

> **Jack and Jill**. Jack believes that whatever Jill says next will be false. In a nearby world, the next thing Jill says is "whatever Jack believes is true".

Intuitively, Jack doesn't know that whatever Jill says next will be false, because in a nearby world that belief is not true – although that belief is not obviously false. Again, this intuition favours (BELIEF-SAFETY) over (VOGEL-SAFETY).

Let's assume that (BELIEF-SAFETY) is correct. Now recall:

> **(MAR)** If K(C) obtains at α_i then C obtains at α_{i+1}

To derive (MAR) from (BELIEF-SAFETY), what is needed is a principle that connects S's belief about C in α_i to S's belief about C in α_{i+1}. After all, asks Berker, might it not be possible for S to stop believing that she feels cold the precise moment she no longer feels cold? He writes:

> [W]ho is to say that…as one gradually gets warmer and warmer during the course of the morning while carefully attending to how cold one feels, one stops feeling cold *before* one stops believing that one feels cold? (2008, p. 8).

If so, Berker suggests, then S's belief that she feels cold would satisfy (BELIEF-SAFETY) without (MAR)'s being true. To close the gap between (BELIEF-SAFETY) and (MAR), Berker proposes the following:

---

[13] For a discussion of how best to formulate safety in light of this sort of consideration, see Manley (2007).

[14] I'm assuming that the utterance "that's a tollbooth" expresses a belief in both the actual and nearby cases. This is not consistent will all accounts of singular thought.

(BEL) If S believes C obtains in α_i, then S believes C obtains in α_{i+1}

And indeed, (MAR) follows easily from (BELIEF-SAFETY) and (BEL). If S's believing she feels cold at one moment entails that she believes she feels cold at the next, then by (BELIEF-SAFETY) if S *knows* she feels cold at one moment she must feel cold at the next. But (BEL) is untenable.[15] It is a soritical premise, since one can generate a paradox from it along with the assumptions that S believes that she feels cold at dawn and that S does not believe that she feels cold at noon (Berker (2008), 7; Vogel (2010), 561). By the description of Cold Morning, S believes she feels cold at dawn and stops believing she feels cold sometime later; it cannot be true that belief in one case entails identical belief in the next. (BEL) is quite obviously not always true.

No matter. Nothing as strong as (BEL) is needed to derive (MAR) from (BELIEF-SAFETY). According to (BELIEF-SAFETY), knowledge that one is cold is incompatible with untrue belief that one is cold in sufficiently similar cases. Now, imagine that in α_i S feels cold, and in α_{i+1} it is no longer the case that she feels cold. And imagine that in α_i she believes truly that she feels cold and in α_{i+1} she doesn't believe she feels cold. That is, her belief that she feels cold stops immediately with the cessation of her feeling cold, as per Berker's suggestion. Does S's belief that she feels cold in α_i satisfy (BELIEF-SAFETY)? Not necessarily. For it could well be that in some sufficiently similar non-actual case β_{i+1}, S continues to believe she feels cold after she stops feeling cold. α_i is a case in the actual world, as is α_{i+1}. But there are also *non*-actual cases that are sufficiently similar to α_i to destroy knowledge if S untruly believes in them that she is cold. To pass the safety test for knowledge, it is insufficient that one, as a matter of chance, lack untrue belief in all actual similar cases. One must also lack untrue belief in *possible* similar cases. This means that it is much easier for S to fail to know that she is cold than some luminists seem to think. All that is required to derive (MAR) from (BELIEF-SAFETY) is something like the following.[16]

> **(BEL*)** If in case α_i S believes she feels cold, then there exists a sufficiently similar possible case β_{i+1} in which S's cold-feelings are a phenomenal duplicate of her cold-feelings in α_{i+1} and in which S believes she feels cold.

[15] That is, understood as a universal generalisation. This applies to my discussion of (BEL*) as well.

[16] Berker suggests, but then rejects, a similar modalised version of (BEL), which he calls (BEL') (2008, p. 7, n. 11). I address Berker's objection to (BEL*)/(BEL') shortly.

Together, (BELIEF-SAFETY) and (BEL*) yield (MAR). If in a given actual case, S knows that she feels cold, then by (BELIEF-SAFETY) there cannot be any very similar cases in which she believes she is cold but isn't. According to (BEL*), if S believes she feels cold at one moment, there is some nearby world in which she believes she feels cold in the next moment (holding her feelings of cold at that moment fixed) (cf. Vogel (2010), 562). So, if in a given actual case S knows she feels cold, then she must feel cold in the next actual case – viz., (MAR).

(BEL*), unlike (BEL), is highly plausible.[17] First, (BEL*) is just the kind of thing you would expect to be true of creatures like us. This is because we don't just believe at random. Our mental lives are structured by certain dispositions. When we believe something in one set of circumstances, in very similar circumstances we have a disposition to believe the same thing. (BEL*) should be understood as encoding the empirical assumption that S, being a creature like us, shares these dispositions.[18,19] We might call this the *doxastic disposition premise*:

> **(DOXDIS)**: If in condition R, S believes she is F, then for any condition R' very similar to R, S has some disposition in R' to believe she is F.

---

17 Again, on the assumption that there is no constitutive connection between the phenomenal and the doxastic.

18 Cohen (2010, 727-30) discusses deriving (MAR) from an empirical premise and a safety condition on knowledge (though Cohen's empirical premise repeats Williamson's talk of 'indiscriminability', which I find unhelpfully opaque). Cohen ultimately argues that this strategy does not work because it makes Williamson's argument circular:

> [According to Williamson] our judgment that α_{i+1} is similar to α_i may require the judgment that if one could wrongly believe one feels cold in α_{i+1}, then in α_i one does not know one feels cold. Given [the empirical premise], this requires the judgment that in α_i one knows one feels cold only if in α_{i+1} one feels cold. And this is just the judgment that [(MAR)] is true (729) (cf. Blackson (2010, 402)).

Quite. Williamson's anti-luminosity argument requires that one make certain judgments about similarity, and on Williamson's view, these judgments are interdependent with our judgments about knowledge. Thus the anti-luminosity argument shows that if one wants to defend luminosity, one must deny that the cases in Cold Morning (and the possible phenomenologically identical cases) *are* similar, which seems absurd. This might be circular, but it is only *viciously* so if the luminist were antecedently willing to deny what seems (to me at least) evidently true.

19 There might be certain cases of possessing dispositions to believe p that do not entail having a nearby belief in p: for example, the disposition to believe of the Müeller-Lyer illusion that the lines are of different length. Since I plausibly retain such a disposition even when it is well suppressed by my awareness of the illusion, this might be a case where a disposition to believe p doesn't entail a nearby belief in p. But such unusual cases are not my concern here.

(DOXDIS) seems to me fairly uncontroversial.[20] Imagine I am looking at a jar full of 1000 marbles; I don't know how many marbles there are in the jar, but I form the belief that there are lot of marbles in the jar. If I'm then confronted with a very similar scenario – a jar with, say, 999 marbles, at a similar distance and in similar lighting conditions, etc. – I am disposed in that scenario to believe, again, that there are a lot of marbles in the jar. This is so even if I don't as it happens believe, in the second scenario, that there are a lot of marbles in the jar. Of course, what counts as a 'very similar condition' in (DOXDIS) will turn on, in part, what agents have the disposition to believe in the relevant situations. If I totally lack the disposition to believe the same in two situations,[21] this in part constitutes their not being very similar.

If this empirical characterisations of our doxastic dispositions is correct, then the luminist who wants to accept (BELIEF-SAFETY) while rejecting (MAR) is under pressure to deny a highly intuitive picture of how we work. In particular, he will have to maintain that, in the first case of not feeling cold, S has no disposition whatsoever to believe that she feels cold. It is insufficient, *pace* various luminists, that S simply happen not to believe that she feels cold when she stops feeling cold; she must lack even the disposition to so believe. But it is implausible that S – since she is, *ex hypothesi*, a creature like us – lacks such a disposition.[22] Thus it is implausible that any conditions of interest are, for us, luminous. Second, (BEL*), unlike (BEL), is not obviously a soritical premise. It does not trade on the vagueness of 'believes', but is instead a specific claim about what is true of S in Cold Morning; our assent to is secured by what we know about the doxastic dispositions of creatures like ourselves in situations like Cold Morning.

While Berker acknowledges that (BEL*) is not straightforwardly soritical in the way that (BEL) is, he suggests that (BEL*) nonetheless has a soritical consequence. This is because repeated applications of (BEL*) yield the conclusion that it is possible for S to feel extremely hot while believing she

[20] Though it might demand some refinement. Imagine that there is an extremely low objective chance that Hanna will believe p in condition R. Despite the terrible odds, she comes to believe p in R. Does she have a disposition to believe p in circumstances that are extremely similar to R? Perhaps not. But the issue needn't worry us here; I mean (DOXDIS) only to be able to handle the kind of central cases of belief-dispositions at issue here.

[21] Here and throughout I talk about having 'the disposition to believe the same in similar situations'. By this I mean not that for any two similar situations we will have the same disposition to believe in those situations, but rather that if in the first situation we believe something, then in the second we have a disposition to believe the same thing.

[22] Some luminists appear to think that what is at stake dialectically in the anti-luminosity debate is whether there are any possible creatures who enjoy luminosity, not whether we are such creatures ourselves. See section 5 for a discussion of why this is misguided.

feels cold.[23] Berker claims that this is an unacceptably absurd consequence of (BEL*).[24] He writes:

> I think we should have serious doubts that such a case is even possible – serious doubts that there could exist a being who counts as having beliefs and experiences, and yet whose beliefs and experiences are as wildly at odds with one another as they would be in [the case in which one feels extremely hot but believes oneself to feel cold] (2008, n. 11, 7-8).

Is it really so hard to countenance such a possibility? The similarity relation is intransitive, so a case in which one felt extremely hot but believed oneself to feel cold would be a case very dissimilar to the one imagined in Cold Morning. In particular, the intransitivity of 'very similar method' means that, in such a case, one might very well be using a method very different from the one a normally functioning person uses to form beliefs about her feelings of cold. One could, for example, be the victim of prolonged psychological priming, or in the grip of a certain philosophical picture of the mind that makes one systematically distrust one's inclinations to judge one's own phenomenal state. Is it really so hard to imagine someone in these conditions coming to believe she feels cold when she actually feels extremely hot? These possible cases might be remote, no doubt. But their existence – like the existence of bad sceptical worlds – does nothing to undermine S's ability to know in normal situations. That (BEL*) implies that they are possible is thus no knock against it. In any case, as Berker himself notes, this objection to (BEL*) seems motivated by a view on which the phenomenal and the doxastic enjoy a constitutive connection. Such a view is not my target here, and (BEL*) will not feature in my argument against it.

Before moving to my second defence of (MAR), one that ~~loosens~~ the assumption that there is no constitutive connection between the phenomenal and the doxastic, let me take quick stock. On my favoured reconstruction of the anti-luminosity argument (for non-constitutive accounts), its essence is this. Imagine that S, like us, is the sort of creature for whom believing something in one situation means having, in extremely similar

---

[23] Berker in fact claims that repeated applications of (BEL*) yield the conclusion that one could believe oneself to feel cold while feeling as hot as "if one were in the center of the sun" (n. 11, p. 7). Since it's doubtful that one would feel much of anything if one were in the centre of the sun, I take it that Berker just means 'extremely hot'. It's worth noting that one can't generate this consequence from repeated applications of (BEL*) alone – one would need analogous principles that apply not just to α cases, but also β cases and so forth. It's also worth noting that ~~if, as seems not implausible,~~ there is an upper bound to how cold or hot one can feel, then not all of these analogous principles will be true.

[24] Thanks to an anonymous referee for pointing this out to me.

situations, the disposition to believe the same thing. Suppose that S is in a condition C, but in what we might call a 'liminal' case of it.[25] That is, there is an extremely similar case to the one she is in which is not a case of C. Imagine that S believes she is in C; is this belief knowledge? It seems not. For in the very similar non-C case, S has the disposition to believe she is in C. Thus, her true belief that she is in C is rendered unreliable by a nearby untrue belief that she is in C. Thus in such a liminal case of being in C, S cannot know that she is in C. This argument generalises to all non-trivial conditions and for all subjects whose beliefs are structured by these sorts of dispositions. That is, it generalises to all interesting mental state conditions in which creatures like us plausibly find ourselves.

My reconstruction of the anti-luminosity argument differs from the original in avoiding Williamson's favoured talk of our 'limited powers of discrimination'. According to Williamson, it is this cognitive limitation that drives the anti-luminosity argument (2000, 12, 97, 103-4). While Williamson unpacks the idea in various ways,[26,27] I take its essence to be this. When we are thinking about whether or not we are in some sort of state, we turn our attention to the relevant underlying phenomenon that constitutes that state. When it comes to figuring out whether there is any milk in the fridge, we train our attention on the contents of the fridge. Similarly, when it comes to figuring out whether we feel cold, we turn our attention to our

[25] The appeal to liminal cases here does not appeal to the vagueness of the concept of C. One can imagine those borders firmly fixed. So, for example, imagine that S is in condition C just in case she is experiencing at least 100 units of x (where x is some phenomenal experience, e.g. feelings of cold), and imagine that she is experiencing exactly 100 units of x. I avoid the more obvious terminology of 'borderline' cases because of its associations with semantic/conceptual vagueness.

[26] Some of which can sound somewhat question-begging. Consider the following:

> The main idea behind the argument against luminosity is that our powers of discrimination are limited. If we are in a case α, and a case α' is close enough to α, then for all we know we are in α'. Thus what we are in a position to know in α is still true in α'. Consequently, a luminous condition obtains in α only if it also obtains in α', for it obtains in α only if we are in a position to know that it obtains in α. (2000, 12).

One might have a similar worry about Williamson's stipulation that "[S] is not aware of any change in [her feelings of cold or hot] over one millisecond" (2000, 97) and that she is "almost equally confident that [she] feels cold, by the description of the case" (2000, 97). These glosses on the argument can sound more like re-descriptions of its conclusion than reason to accept it.

[27] Cf. Vogel (2010, part II) for a discussion of what Williamson might mean by 'limited powers of discrimination', and how this generally relates to the luminosity doctrine.

sensations of cold.[28] Now, it is a disappointing truth of our perceptual capacities that they are not infinitely discriminating: we cannot always tell of two distinct things whether they are indeed distinct. If, by chance, we come to believe that two indiscriminable situations are in fact different, this lucky belief does not constitute knowledge. Thus, our limited perceptual capacities limit our ability to know. Williamson's anti-luminosity argument can be understood as an application of this observation to the phenomenal sphere. In training our attention on our underlying sensations of cold and hot, we cannot reliably distinguish between two extremely similar sensations, one of which is cold and the other of which is not. Since reliability is required for knowledge, we cannot know that we are cold in such a case.[29] Of course, the luminist can simply reply that this too begs the question against him. For any defender of the luminous could simply reject the idea that our powers of discrimination *are* limited. But this is to maintain not only that we have privileged access to our mental states, but also that our perceptual capacities when it comes to attending to those mental states, unlike attending to external world states, are infinitely discriminating. And this might seem like a large bullet to bite.

How does this interpretation of Williamson's original argument square with my own reconstruction? My empirical premise – (DOXDIS) – is certainly compatible with Williamson's (also empirical) claim that our powers of discrimination are limited. But it is also compatible with a variety of other stories we might want to tell about how our phenomenal beliefs arise.[30] Whatever version of that story we embrace, it seems undeniable that creatures like us our disposed to believe the same things in extremely similar situations. Of course, what we mean by 'similar' here matters crucially. The luminist can always resist the claim, necessary for (BEL*) to follow from (DOXDIS), that the situations from one moment to the next in Cold Morning are, indeed, 'very similar'. And, as Williamson himself admits, our judgments about similarity are inevitably bound up with our epistemic judgments

[28] This presupposes a broadly perceptual model of self-knowledge. For a discussion of alternative models and how they interact with the anti-luminosity argument, see section 5.

[29] We might worry that this argument isn't enough. The mere ability to distinguish dissimilar things is not the same thing as the ability to correctly categorise them under concepts. Thus I might be able to distinguish two distinct phenomenal sensations without being in a position to know that one is a feeling of cold and another isn't. (Compare: I might be able to distinguish between the sound of French and the sound of German, but this doesn't mean that I'm always in a position to know that I'm hearing French rather than German). If so, then this is another reason to prefer my version of the anti-luminosity argument in terms of dispositions to believe.

[30] These alternative stories become particularly important when we consider the possibility, discussed in the next section, that the phenomenal is constitutively connected to the doxastic.

(Williamson 2000, 100-1). So the luminist can always dig in his heels somewhere; my formulation of the argument in terms of doxastic dispositions will not compel the dogmatic luminist to abandon his views. But, if my reconstruction of Williamson's argument works, then the price the luminist has to pay for resisting it is very high indeed. He must either deny that we are creatures who are disposed to believe ~~similarly~~, in similar situations, or insist that the cases from one moment to the next in Williamson's thought experiment should not count as similar. Both options seem to me perverse. As such, defending (MAR) by invoking a safety condition on knowledge together with the minimal (DOXDIS) seems to me dialectically preferable to Williamson's own defence in terms of finite powers of discrimination.

5. Defending (MAR): Constitutive Accounts

In defending (MAR) against the luminists, I argued that it derived easily from a simple safety condition on knowledge, (BELIEF-SAFETY), coupled with the plausible premise (BEL*), which in turn is justified by the empirical premise (DOXDIS). However, (BEL*) is subject to complaint from those luminists who maintain that phenomenal conditions such as feeling cold have a constitutive connection to one's beliefs about whether one feels cold. As such, the anti-luminosity argument as elaborated so far is impotent against a popular view of the phenomenal (Weatherson (2004), Berker (2008) and Ramachandran (2009)).[31] Imagine that the following were true of the relationship between S's beliefs about feeling cold and the facts about her feeling cold:

> **(CON)** If S has done everything she can to decide whether she feels cold, then she believes that she feels cold if and only if she feels cold.[32]

Since by the description of Cold Morning, S is doing everything she can to decide whether she feels cold, her coming to believe that she feels cold is both necessary and sufficient for her indeed feeling cold. Now recall (BEL*):

> **(BEL*)** If in α_i S believes she feels cold, then there exists a sufficiently similar possible case β_{i+1} in which S's cold-feelings are a phenomenal duplicate of her cold-feelings in α_{i+1} and in which S believes she feels cold.

---

[31] Berker and Ramachandran both propose that feeling cold might be a response-dependent condition, while Weatherson argues along physicalist lines that feeling cold and believing oneself to be cold could in fact consist in the same brain state. There are various possible ways of unpacking what the constitutive connection between feeling cold and believing oneself to feel cold might amount to.

[32] Adapted from Berker (2008), p. 9.

If (CON) is true, then (BEL*) is false. Why? By the description of Cold Morning, there is some value i such that S believes she feels cold in α_i and no longer believes she feels cold in α_{i+1}. (CON) entails that for that value of i, S does indeed feel cold at α_i and no longer feels cold α_{i+1}. And according to (CON), any case that is a phenomenological duplicate of α_{i+1} (with regard to S's feelings of cold) will also be a *doxastic* duplicate of α_{i+1}. So (BEL*) goes false for the value of i such that S believes she feels cold in α_i and no longer believes she feels cold in α_{i+1}. For that value of i, it is not the true that there exists a possible case β_{i+1} in which S's cold-feelings are a phenomenal duplicate of her cold feelings in α_{i+1} and in which S believes C obtains. Thus, if (CON) is true, (BEL*) is false.

To defend (MAR) without assuming, as we have been doing, that (CON) is false, we need to appeal to Williamson's refined safety requirement in terms of degrees of confidence.[33,34] We might specify this safety condition as follows:

> **(CONFIDENCE-SAFETY)** If in case α S knows with degree of confidence c that she is in a condition R, then in any sufficiently similar case α* in which S has an at-most-slightly-lower degree of confidence c* that she is in condition R, it is true that she is in condition R.

The idea behind (CONFIDENCE-SAFETY) is that for one to know that one is in a given condition, it cannot be the case that one is almost as confident that one is in that condition – *even if* that confidence is short of full-fledged belief – in a sufficiently similar situation. That is, nearby misplaced

---

[33] Williamson distinguishes these from degrees of subjective probability measured by one's betting behaviour. A degree of confidence is a degree of *outright belief*:

> Intuitively, one believes p outright when one is willing to use p as a premise in practical reasoning. Thus one may assign p a high subjective probability without believing p outright, if the corresponding premise in one's practical reasoning is just that p is highly probable on one's evidence, not p itself…we can think of one's degree of outright belief in p as the degree to which one relies on p. Outright belief in a false proposition makes for unreliability because it is reliance on a falsehood (2000, p. 99).

On Williamson's view, one can have a certain degree of confidence (outright belief) without having *an* outright belief. This non-standard distinction has created a lot of confusion about Williamson's argument (e.g. Blackson 2007). Cf. Ramachandran's discussion (2009, 663, especially footnote 3). Ramachandran's first interpretation (he offers four in total) of Williamson's argument relies on a (deliberate) misreading of Williamson's notion of confidence.

[34] Cohen (2010) argues, quite correctly, that (CONFIDENCE-SAFETY) cannot be derived from (BELIEF-SAFETY) without a soritical premise, namely that a slight change in degree of confidence does not affect whether one believes outright. He thus concludes that (CONFIDENCE-SAFETY) cannot be generally motivated. However, Williamson intends (CONFIDENCE-SAFETY) to be his fully elaborated safety condition, itself an intuitively plausible gloss on the reliability requirement for knowledge. Here I defend it as such.

confidence – high confidence in an untruth – is sufficient to preclude knowledge. Let us grant for the sake of argument that S's feeling cold is constitutively connected to her belief that she feels cold in the way (CON) specifies. ~~Imagine~~ that in α_i S truly believes that she feels cold, and that in α_{i+1} she is still quite confident that she feels cold, but insufficiently confident for outright belief. By (CON), S feels cold in α_i but does not feel cold in α_{i+1}. But by (CONFIDENCE-SAFETY), S does not know that she feels cold in α_i. So even a constitutive connection between feeling cold and believing one feels cold is insufficient to vindicate luminosity.

Again, to derive (MAR) from (CONFIDENCE-SAFETY) one needs an additional assumption that links S's confidence that she feels cold to her confidence that she feels cold in nearby cases. Berker proposes something like the following:

> **(CONF)** If in α_i S has degree of confidence c that she feels cold, then in α_{i+1} S has an at-most-slightly-lower degree of confidence c^* that she feels cold

While Berker does not wish to dispute (CONF) – saying that it "seems indisputable, given the description of the situation at hand" (2008, p. 12) – it is again worth noting that, as with (BEL), (CONF) is unnecessarily strong. To generate (MAR) from (CONFIDENCE-SAFETY), the following weaker premise will do:

> **(CONF*)** If in α_i S has degree of confidence c that she feels cold, there exists a sufficiently similar possible case β_{i+1} in which S's cold-feelings are a phenomenal duplicate of her cold-feelings in α_{i+1} and in which S has an at-most-slightly-lower degree of confidence c^* that she feels cold

It should be clear that (CONF*) is even more plausible than (CONF), again by appeal to an empirical premise:

> **(DOXDIS*)** If in condition R, S believes with confidence level x that she is F, then for any condition R' very similar to R, S has some disposition in R' to believe with confidence levels similar to x that she is F.

Roughly, (DOXDIS*) is the claim that if in a certain situation S has a certain confidence level, then in a very similar situation she is disposed to have a very similar confidence level.[35] As such, if at α_i S has degree of confidence c that she feels cold, there exists a sufficiently similar possible case in which S feels just as cold as she does in α_{i+1} and has a degree of

[35] (DOXDIS) and (DOXDIS*) can be subsumed under a more general empirical principle: If in condition R S has attitude A towards p, then for any similar attitude A' and similar condition R', S in R' has some disposition to have A' toward p.

confidence c^* that she feels cold that is at most slightly slightly lower than c (and thus similar to c). Like (BEL*), (CONF*) is not a soritical premise, since it cannot be used to generate the (obviously false) conclusion that S has the same or similar degree of confidence at dawn and noon that she is in C. Rather than trading on the vagueness of any of its constituent terms, (CONF*) encodes a plausible empirical premise about our dispositions to believe similarly in similar situation. And finally, unlike (BEL*), (CONF*) is compatible with (CON). While (BEL*) entailed that S's belief about whether she feels cold could possibly come apart from the fact of whether she feels cold – in contradiction with (CON) – (CONF*) doesn't entail anything of the sort. Instead, (CONF*) merely entails that S's confidence levels about her feelings of cold are similar in similar cases. This is perfectly compatible with the response-dependent view that S's believing she feels cold is both necessary and sufficient for her feeling cold.

In any case, it is at (CONFIDENCE-SAFETY) that many luminists direct their attack (Leitgeb (2002), Berker (2008), Ramachandran (2009), Cohen (2010)). They hope to show that (CONFIDENCE-SAFETY) is a non-genuine condition on knowledge, leaving us only with the original defence of (MAR) in terms of (BELIEF-SAFETY), and thus the constitutive connection view of the phenomenal unscathed by the anti-luminosity argument. For example, in his attack on (CONFIDENCE-SAFETY), Cohen argues:

> [I]t is not obvious why one's confidence at t_i is misplaced. We are supposing that at t_i one knows one feels cold. Thus at t_i one feels cold and one believes one feels cold. It follows that if at t_{i+1} one no longer believes one feels cold, then at t_i one just barely believes one feels cold. Now suppose one no longer feels cold at t_{i+1}. Then one just barely feels cold at t_i. So under these suppositions, at t_i one just barely feels cold and one just barely believes one feels cold. So how is one's confidence at t_i misplaced? (Cohen 2010, 726).

In a similar spirit, Berker writes:

> [W]hy should we withhold the honorific 'reliable' in the kinds of cases Williamsons describes? …[W]hat if one's belief that p tapers off (as it were) just as its being the case that p tapers off, and in precisely the same way?…(CONFIDENCE-SAFETY) deems as unreliable belief-forming mechanisms that appear to be as reliable as they could possibly be (2008, p. 12).

Spelling out this line of objection, Berker proposes that we think of S's feelings of cold in terms of numbers of "freezons". He imagines that at dawn S experiences 50 freezons worth of cold, that at noon she experiences -50 freezons worth of cold, and at any time during the day her degree of

confidence that she feels cold is directly correlated with her subjective feeling of cold as measured in freezons. Finally, Berker supposes that S believes she feels cold if and only if she indeed feels cold. That is, S's confidence that she feels cold drops below the belief-threshold at precisely the same moment she ceases feeling cold. If (CONFIDENCE-SAFETY) is true, then S, at some point during the day, fails to know that she feels cold. But Berker objects that this is absurd, since "one's beliefs about whether one feels cold appear to be as reliable as they possibly could be" (2008, p. 13). In the same vein, Ramachandran writes:

> [(CONFIDENCE-SAFETY)] is too strong a requirement because it would rule out luminosity in the hypothesized 'perfect-calibration' situation, which is daft, because one *couldn't* be any more reliable (2009, 668; cf. Leitgeb (2002), 216).[36]

While there is something no doubt attractive in the Leitgeb-Berker-Ramachandran-Cohen thought that this 'perfect calibration situation' represents some sort of maximally reliable possibility – and thus that (CONFIDENCE-SAFETY) can't possibly be a genuine condition on knowledge – that attraction isn't too difficult to shake. Consider the following case, analogous to Berker's freezons case:

> **Glass Half Full**. Henry likes watching empty glasses slowly fill with water until they are full. In normal conditions and when he is paying close attention, Henry's confidence that a given glass is at least hall full is directly correlated with how full the glass is, rising slowly from 0% to 100% confidence as the initially empty glass is slowly filled to the brim. Moreover, Henry believes that glasses are at least half full if and only if they are indeed at least half full. It thus follows that the confidence threshold for outright belief is 50%. The only proposition Henry entertains about a glass as it fills is *that the glass is at least half full*.

Does Glass Half Full, as Berker et al. suggest, represent a 'perfect calibration situation'? Surely not. When the glass is only a fifth full – that is, very obviously less than half full, even to an average estimator – Henry still has a 20% confidence that the glass is at least half full. And when the glass is four-fifths full – that is, very obviously more than half full, even to an average estimator – Henry has only an 80% confidence that the glass is at least half full. Henry's confidence profile is hardly a maximally reliable possibility. The suggestion that Henry, or Berker's freezons subject, represents a

[36] Ramachandran (668) ultimately does not endorse this line of argument on the grounds that the individual it imagines is too different from the subject in Cold Morning. I take this to be a bad objection because the issue is whether (CONFIDENCE-SAFETY) is a genuine condition on knowledge in general.

'perfect calibration situation' is thus misguided. So it cannot be reason to think that either case constitutes a counterexample to (CONFIDENCE-SAFETY).

For a more plausible counterexample to (CONFIDENCE-SAFETY), take the following case:

> **Glass Half Empty**. Henrietta likes watching full glasses slowly drain until they are empty. The confidence threshold for outright belief is 80%, and, in normal conditions and when she is paying close attention, Henrietta believes of a given glass that it is at least half empty if and only if it is indeed at least half empty. Henrietta's confidence that the glass is at least half empty remains near 0% until the glass is almost half empty, then steeply increases to 80% just as the glass is exactly half empty, increasing to and remaining at 100% shortly thereafter. The only proposition that Henrietta entertains as the glass empties is *q*, *that the glass is at least half empty*.

Henrietta, unlike Henry, is a very good judge of the fullness or emptiness of glasses. For Henrietta doesn't have relatively high confidences in *q* when it is obviously false (e.g. when the glass is only a fifth empty), or relatively low confidences in *q* when it is obviously true (e.g. when the glass is four-fifths empty). Since all of Henrietta's *q*-beliefs are reliably true, they satisfy (BELIEF-SAFETY).[37] Nonetheless, not all of Henrietta's *q*-beliefs satisfy (CONFIDENCE-SAFETY). Specifically, Henrietta's belief in *q* when it is exactly half empty does not satisfy it. For just a moment earlier – when the glass was *almost* half empty – Henrietta had a high confidence just short of belief that the glass was at least half empty. If it seems plausible that Henrietta knows that the glass is at least half empty when it is exactly half empty, this tells against (CONFIDENCE-SAFETY) in favour of the less demanding (BELIEF-SAFETY).

The crucial question then is: does Henrietta know that the glass is at least half empty when the glass is exactly half empty? There is reason to think not. After all, even if Henrietta reliably tips over the threshold to outright belief just when it becomes true that the glass is at least half empty, it's also the case that Henrietta reliably *almost* believes that the glass is at least half empty when that's false. That is, in nearby cases in which it is false that the glass is at least half empty, Henrietta has a high confidence that it is true. Specifically, at the point at which the glass is just less than half empty, Henrietta is 79% confident that the glass is at least half empty: very

[37] So do Henry's glass-beliefs, of course, since he believes a glass to be at least half full if and only if it is indeed so. So what I say about Henrietta goes for Henry as well; I've only introduced the Glass Half Empty case to clear away the distractions that arise from Henry's putatively 'perfect' calibration.

confident indeed. Does this not undermine Henrietta's claim to know q just a moment later, when the glass is exactly half empty?

A common picture of the relationship between confidence and practical reasoning would suggest so. According to that picture, one's confidence in a proposition p is a measure of one's willingness to rely on p as a premise in practical reasoning; one believes outright when one's willingness to rely on p crosses a certain threshold.[38] If so, then a confidence just short of outright belief in p will yield some cases in which one uses p in one's practical reasoning despite not believing p outright. When the glass is only, say, 49% empty, Henrietta's 79% confidence that the glass is at least half empty means that she has some tendency to use that false proposition as a premise in her practical decision-making. If she were making several q-relevant but independent decisions at the same time, we could expect to see Henrietta acting on q when it is false. This nearby willingness to act on q when it is untrue intuitively undermines Henrietta's claim to know q in the case where q has just become true.

Perhaps though the most compelling case in favour of (CONFIDENCE-SAFETY) can be made by tweaking the thought experiments that motivated safety in the first place. Take for example:

> **Receding Fake Barns**. Mirra is looking at two rows of what look like barns in the distance. The first row is made up of real barns; the second row is fake. In situations like this, Mirra only forms beliefs about the proposition *that is a row of barns*, and she reliably forms only true beliefs about that proposition. The threshold for outright belief is 70% confidence. Of the first row, Mirra believes with 70% confidence that it is a row of barns. Of the second row, Mirra believes with 69% confidence that it is a row of barns.

Mirra's belief that the first row is a row of barns is reliably true; she could not have easily had the untrue belief that it is a row of barns. But it nonetheless seems somewhat odd to say that Mirra knows that the first row is a row of barns. After all, Mirra has a 69% confidence that the fake barns right behind the real barns constitute a row of barns. This suggests that safety requires more than the absence of nearby untrue belief; it requires the absence of nearby untrue *almost*-belief. If so, then luminosity requires not only that our phenomenal beliefs satisfy (BELIEF-SAFETY); it requires that our phenomenal beliefs further satisfy (CONFIDENCE-SAFETY). Since the former but not the latter can be satisfied by beliefs that enjoy a constitutive connection to their phenomenal contents, it seems that no non-trivial mental conditions are luminous.

[38] That is to say, one's believing p isn't anything over and above meeting a certain threshold for confidence – e.g. putting p in one's belief-box or mentally asserting p.

This is not to say that *no* possible creature could have a degree-of-confidence profile that satisfied both (CONFIDENCE-SAFETY) and luminosity. Take again Henrietta, who believes q, that the glass is at least half empty, if and only if q is true. Imagine that at the point at which q becomes true, Henrietta's confidence level discontinuously jumps from 0% to 80%. In so doing, Henrietta believes not just truly but moreover *safely* that the glass is at least half empty whenever it is indeed at least half empty. Thus for her the condition of the glass being at least half empty is luminous. Analogously, if the subject of Berker's freezon case were to exhibit a jump down from a high to low confidence at the point at which she no longer feels cold,[39] then for her, feeling cold would be luminous. As Berker himself notes, physical systems are rarely characterised by such discontinuity (2008, 15). As such, it's very implausible for creatures like us anything very much is luminous.

Berker however draws a different lesson. He claims that to maintain that (CONFIDENCE-SAFETY) is a genuine condition on knowledge is to "insist not just that *reliability* is required for knowledge, but moreover that *perfect reliability* is required, and that way scepticism lies" (ibid., 15). But this is misleading. It is the objector to (CONFIDENCE-SAFETY), not its proponent, who demands perfection. Even Henry, whom all will admit is a pretty dismal estimator, gets to know some of the time, according to (CONFIDENCE-SAFETY), that the glass is at least half full. And similarly for Berker's freezon subject. (CONFIDENCE-SAFETY) will deliver even more generous verdicts when it comes to imperfect epistemic agents like us; while diverging quite a bit from the maximally reliable confidence profile, we still get to know much of the time that we are in various conditions. (CONFIDENCE-SAFETY) doesn't generally demand perfection: far from it. Rather, it implies that we are not perfect, which is a different thing altogether. Our implied imperfection is no reason to reject (CONFIDENCE-SAFETY). Or, rather, it is no reason beyond an antecedent affection for luminosity.

It's worth noting that the luminist could simply deny that (CONFIDENCE-SAFETY) is a genuine condition on knowledge of the phenomenal, even if he accepts it as a genuine condition on knowledge of external states like the fullness of a glass or the presence of barns. His willingness to do so will turn on the particular picture of the phenomenal-doxastic connection

[39] Berker describes such a confidence profile (see his Figure 2, p. 14) as one in which one is absolutely certain (i.e. has confidence 1) that one feels cold when one indeed feels cold, and drops off to confidence 0 when one doesn't feel cold. But nothing this extreme is required. (CONFIDENCE-SAFETY) and the luminosity of feeling cold are together consistent with a more leisurely descent in confidence both above and below the belief-threshold; all that is required is a discontinuous drop-off at the point where one ceases to feel cold (in the actual and very similar cases).

to which he is committed, or willing to commit himself in order to vindicate luminosity. If, for example, one maintained that feeling cold were simply a matter of believing oneself to feel cold, perhaps after deliberation – that there is no underlying phenomenon of 'feeling cold' that our beliefs are tracking or failing to track – then one could plausibly insist that models of reliability that are based on perceptual capacities are here inapplicable. Thought experiments like Glass Half Empty and Receding Fake Barns are thus orthogonal to the issue. In this way, a commitment to (CON) might still offer a weapon of resistance against (MAR), and thus the anti-luminosity argument. But for reasons I discuss in the next section, the weapon might be double-edged.

6. Conclusion: The Dialectical Situation

Let me take stock. Luminists typically argue that (MAR) cannot be derived from plausible premises: either it requires an overly strong safety condition, or it requires additional soritical premises. I've offered a reconstruction of Williamson's two-part defence of (MAR), the first intended to work on the assumption that there is no constitutive connection between the doxastic and the phenomenal, and the second intended to work without ~~even without~~ that assumption. In each case I've argued that an independently plausible safety condition – (BELIEF-SAFETY) and (CONFIDENCE-SAFETY), respectively – combined with an independently plausible view of our empirical characters, yields (MAR). In brief, our coarse-grained dispositions to believe renders us incapable of knowing that we are in liminal cases of (otherwise luminous) conditions. This is plausibly true whether or not there exists a constitutive connection between the phenomenal and the doxastic, at least assuming that such a connection preserves the status of self-knowledge as an instance of knowledge more generally. Thus the luminist who wants to dig in his heels either must reject a plausible picture of our empirical character, or isolate self-knowledge as a sui generis epistemic category. I want to end by saying something about these two possible retreats for the luminist, and how they figure in the overall dialectical situation at hand.

The first possible retreat for the luminist is to reject what I claim is a plausible empirical characterisation of the kinds of creatures we are, namely creatures who have certain coarse-grained dispositions to believe. But some luminists, in arguing against Williamson, seem to think that it is sufficient to show merely that *some* metaphysically possible creature could have a different empirical character, and so enjoy luminosity (Berker 2008, cf. Weatherson 2004, p. 9). Williamson himself is not maximally explicit about the conditions to which his anti-luminosity argument applies: conditions that human and human-like creatures actually find themselves in, conditions that

are metaphysically possible, conditions that are logically possible?[40] But to think the anti-luminosity debate is about possible creatures, not creatures like us, is to mistake the dialectical situation. If the anti-luminosity argument were meant to apply to all possible creatures, we need nothing very much to show it to be impotent. For the anti-luminosity argument is trivially compatible with various possible creatures for whom feeling cold and a whole range of other conditions are luminous.[41] More significantly, the question of whether we are luminous is the question we *should* care about. The possibility of creatures, perhaps radically different from ourselves, for whom interesting conditions are luminous does little to assuage the live possibility that our philosophy of mind, epistemology, and ethics are all built on disreputable Cartesian foundations.

The second possible retreat for the luminist is to endorse a particular kind of constitutive connection view. While a variety of non-perceptual accounts of self-knowledge are compatible with there being a constitutive connection between the phenomenal and the doxastic,[42] according to one such popular account, self-knowledge is deeply disanalogous with knowledge of the external world. That we are always in a position to know our own phenomenal states is a conceptual truth, or a feature of our 'grammar' in the Wittgensteinan sense of that term (Shoemaker 1986; Wright 1989 and 1998; Bilgrami 2006; Heal 2001; Coliva 2009). Self-knowledge is sometimes thus said to be 'no cognitive achievement'. Such a view isolates self-knowledge as a sui generis epistemic state, thereby freeing it from the normal requirements for knowledge like reliability or truth-tracking. Obviously, whether this offers a legitimate retreat for the luminist depends on whether the phenomenal and the doxastic are really connected in just this way – whether self-knowledge is really just a conceptual or grammatical upshot. Again, this is in part an empirical question about the kinds of creatures we are. I don't mean to enter this debate. But I will close with two observations about it.

The first is that those philosophers of mind who endorse constitutive accounts of self-knowledge are often motivated to do so precisely because they wish to vindicate luminosity of the mental. Thus Berker misrepresents the dialectical situation somewhat when he claims that

---

[40] Somewhat gnomically, Williamson writes, prefacing the anti-luminosity argument: "The domain of cases will be taken to include counterfactual as well as actual possibilities. Since the cases on which the arguments below rely are physically and psychologically feasible, issues about the bounds of possibility are not pressing" (2000, p. 94). I take this to mean that Williamson intends Cold Morning to be a description of something that is 'physically and psychologically feasible' for a creature like us.

[41] See n. 10.

[42] See e.g. Chalmers (2003) and various essays in Coliva (2012).

> typically it is *precisely because* they think that there is a tight connection between certain mental states and beliefs about those states that some philosophers claim the mental states in question to be luminous (2008, p. 9).[43]

That some philosophers favour constitutive accounts does not make luminosity more plausible, since it is for the vindication of luminosity that these accounts were generally designed.[44] My second point is this. The tension between Williamson's anti-luminosity argument and the 'no cognitive achievement' view of self-knowledge is deep and fundamental. On Williamson's view, any knowledge worth the name evinces some sort of cognitive achievement. Obviously, fans of the 'no cognitive achievement' view of self-knowledge disagree. But the disagreement here might seem not substantive so much as terminological. In one sense, Williamson has nothing to say to this species of luminist;[45] in another, they have already conceded his central point. To the extent that luminists want to conceive of self-knowledge as an instance of knowledge more generally, the anti-luminosity argument puts pressure on them to abandon luminosity. To the extent that luminists want to preserve luminosity, they are under pressure to accept a different picture of what knowledge itself is.

Acknowledgments

Thanks to Jane Friedman, Jeremy Goodman, John Hawthorne, Clayton Littlejohn, Daniel Rothschild and Timothy Williamson for very helpful discussion of these issues. Special thanks to an anonymous reviewer for detailed and insightful comments on earlier drafts.

References

Berker, S. (2008). Luminosity Regained. *Philosophers' Imprint*, 8(2), 1–22.

Bilgrami, A. (2006). *Self-Knowledge and Resentment*. Cambridge, Mass.: Harvard University Press.

---

[43] Cf. Ramachandran (666).

[44] Of course it might turn out that these accounts, initially motivated by the appeal of luminosity, then turn out to have other theoretical virtues (elegance, fecundity, etc.) and thus support the original datum of luminosity. But then the luminist would need to do more than simply point out that various philosophers endorse a view of the phenomenal as constitutively connected to the doxastic; he would have to show why these accounts recommend themselves independently of luminosity considerations. Thanks to Jane Friedman for raising this point.

[45] ~~I don't mean this literally: Williamson gestures in *Knowledge and its Limits* to other arguments against these kind of views of self-knowledge. But the anti-luminosity argument itself might not be able to get a grip on such views.~~

Blackson, T. (2007). On Williamson's Argument for (I_i) in His Anti-Luminosity Argument. *Philosophy and Phenomenological Research*. Vol. LXXIV No. 2, 397–405.

Brueckner, A. and Fiocco, M. O. (2002). Williamson's Anti-Luminosity Argument. *Philosophical Studies* 110, 285–293.

Chalmers, D. (2003). The Content and Epistemology of Phenomenal Belief. Published in Q. Smith and A. Jokic (eds), *Consciousness: New Philosophical Perspectives*. Oxford: Oxford University Press.

Cohen, S. (2010). Luminosity, Reliability, and the Sorites. *Philosophy and Phenomenological Research* Vol. LXXXI No. 3, 718–30.

Coliva, A. (2009). Self-knowledge and commitments. *Synthese* 171(3), 365–75.

—— (2012). ~~One Variety of Self-Knowledge: Constituitivism as Constructivism. In Coliva A (Ed),~~ *Self and Self-Knowledge*. Oxford: Oxford University Press.

Comesaña, J. (2005). Unsafe Knowledge. *Synthese* 146, 395–404.

Conee, E. and Feldman, R. (1998). The Generality Problem for Reliabilism. *Philosophical Studies*, 89(1), 1–29.

—— (2005). The Comforts of Home. *Philosophy and Phenomenological Research* 70, 444–451.

Hawthorne, J. (2005). Knowledge and Evidence. *Philosophy and Phenomenological Research*, Vol. LXX, No. 2, 452–458.

Heal, J. (2001). On first-person authority. *Proceedings of the Aristotelian Society* CII, 1–19.

Leitgeb, H. (2002). Review of 'Timothy Williamson, *Knowledge and its Limits', Grazer Philosophische Studien* 65, 207–17.

Manley, D. (2007). 'Safety, Content, Apriority, Self-Knowledge' *The Journal of Philosophy*, 104(8): 403–423.

Neta, R. and Rohrbaugh, G. (2004). Luminosity and the Safety of Knowledge. *Pacific Philosophical Quarterly* 85, 396–406.

Ramachandran, M. (2009). Anti-Luminosity: Four Unsuccessful Strategies. *Australasian Journal of Philosophy* 87 (4), 659–673.

Shoemaker, S. (1986). *The First Person Perspective and Other Essays*. Oxford: Oxford University Press.

Vogel, J. (2010). Luminosity and Indiscriminability. *Philosophical Perspectives* 24 (1), 547–572.

Weatherson, B. (2004). Luminous Margins. *Australasian Journal of Philosophy* 82, 373–383.

Williamson, T. (2000). *Knowledge and Its Limits*. Oxford: Oxford University Press.

—— (2005). Replies to Commentators. *Philosophy and Phenomenological Research*. Vol. LXX, No. 2, 468–491.

Wong, W. (2008). What Williamson's Anti-Luminosity Argument Really Is. *Pacific Philosophical Quarterly* 89, 536–543.

Wright, C. (1989). Wittgenstein's later philosophy of mind: Sensation, privacy and intentions. *Journal of Philosophy*, LXXVI, 622–34.

—— (1998). Self-knowledge: the Wittgensteinian legacy. In Wright, C., Smith, B. and Macdonald, C. (Eds.) *Knowing Our Own Minds*. Oxford: Clarendon Press, 13–45.

Zardini, E. (forthcoming). Luminosity and determinacy. *Philosophical Studies*.