Luminosity Guaranteed

Wolfgang Barz

Abstract. This paper aims to show that Williamson's (2000) anti-luminosity argument does not succeed if we presuppose a constitutive connection between the phenomenal and the doxastic. In contrast to other luminists, however, my strategy is not to critically focus on the refined safety condition in terms of degrees of confidence that anti-luminists typically use in this context. Instead, I will argue that, given a certain conception of what Chalmers (2003) calls "direct phenomenal concepts," luminosity is guaranteed even if the refined safety condition in terms of degrees of confidence is taken for granted.

1 Introduction

Williamson's (1996; 2000) anti-luminosity argument is often criticized for neglecting the possibility that there might be a constitutive connection between phenomenal experiences and certain beliefs about them. If there is such a connection, luminists claim, then Williamson's argument is not sound. In response, anti-luminists typically resort to a refined safety condition in terms of degrees of confidence, in light of which phenomenal experiences such as feeling cold do not seem to be luminous, even under the assumption that there is a constitutive connection between the phenomenal and the doxastic. Luminists, in turn, react by criticizing the refined safety condition, arguing that it is too strong a requirement for knowledge. [1] The aim of this paper is to side with the luminists. However, I will not join in the attack on the refined safety condition. Instead, I will argue that, even if the refined safety condition is taken for granted, it does not follow that phenomenal experiences such as feeling cold are not luminous. The key idea is that phenomenal beliefs involving

what Chalmers (2003) calls *direct phenomenal concepts* are immune against epistemic luck.

2 Williamson's argument reshaped

First, let us recall what is at issue: Williamson's (1996; 2000) anti-luminosity argument. Williamson's argument is designed to show that the conditions many think of as paradigmatically luminous, such as being in a certain phenomenal state, are not luminous. Using Williamson's favorite example, it is possible that one fails to know that one currently feels cold, even if (i) one currently does feel cold, (ii) one attends to one's current feeling of cold as assiduously as possible, and (iii) one believes that one currently feels cold on the basis of one's introspective evidence.

Williamson's argument revolves around an example of gradual warming. In order to make the case more vivid, let us look at Brian Weatherson's reformulation:

Mr. Davis's apartment faces southwest, so while it is often cold in the mornings it always warms up as the midday and afternoon sun streams in. This morning Mr. Davis felt cold when he awoke, but now at noon he is quite warm, almost hot. But the change from wake-up time to the present is rather gradual. Mr. Davis does not take a hot bath that morning, nor cook a hot breakfast, but sits reading by the window until the sun does its daily magic. (Weatherson 2004, 375)

Assume furthermore that throughout the period of time described, Mr. Davis carefully attends to how cold he feels. Thus, at every point in time, Mr. Davis holds some belief as to whether or not he currently feels cold. Predictably, Mr. Davis's confidence that he

2

feels cold gradually diminishes: when he woke up, he firmly believed that he felt cold, but now, at noon, he firmly believes that he no longer feels cold but hot. However, note that since the change from feeling cold to feeling hot is extremely gradual and Mr. Davis's powers of discriminations—as with all of us human beings—are limited, he is not able to introspectively distinguish between his thermal experience at some point in time and his thermal experience one millisecond later.

Usually, Williamson's argument is reconstructed as a *reductio ad absurdum* that is driven by a so-called "margin for error" principle to the effect that if Mr. Davis knows that he feels cold at some point in time, then he feels cold one millisecond later. From the assumption that the state of feeling cold is luminous, along with the description of the case, it follows that Mr. Davis knows that he feels cold at t_0, that is, immediately after he woke up. From the fact that Mr. Davis knows that he feels cold at t_0, combined with the margin for error principle, it follows that Mr. Davis feels cold at t_1, which is at an interval of only one millisecond from t_0. Now the luminosity-assumption takes effect again and leads to the claim that Mr. Davis knows that he feels cold at t_1. From this, combined with the margin for error principle, it follows that Mr. Davis feels cold at t_2, which is at an interval of only one millisecond from t_1, and so on and so forth till we arrive at the conclusion that Mr. Davis feels cold at t_n, that is, at noon. Since this is absurd, Williamson argues that the luminosity-assumption must be abandoned.

Although the above reconstruction might be very close to Williamson's original considerations, it is somewhat unfortunate. Williamson's margin for error principle, for example, essentially depends on the idea that safety is necessary for knowledge. Thus, the argument can be straightforwardly rebutted by denying that safety is necessary for

3

knowledge.[2] Moreover, Williamson's margin for error principle has the air of a soritical premise, so that the argument might be dismissed as belonging to the same class of paradoxes such as the Heap or the Bald Man.[3] In my view, both criticisms are beside the point—for Williamson's argument might easily be reformulated without using the aforementioned margin for error principle. For this, we should not perceive it as an argument to the conclusion that it is possible for someone who feels cold to introspect as assiduously as possible without thereby coming to *know* that one feels cold, but rather as an argument to the conclusion that it is possible for someone who feels cold to introspect as assiduously as possible without thereby coming to *safely believe* that one feels cold.[4]

Certainly, the target of the argument, so construed, is considerably weaker than the target of the original version. While Williamson's original argument aimed at refuting the claim that, if one is in a phenomenal state, one is in a position to know that one is in that state, the reshaped argument aims at refuting the claim that, if one is in a phenomenal state, one is in a position to safely believe that one is in that state. However, this weaker claim is nevertheless strong enough to be interesting in philosophical respects as it is an instance of the following scheme:

For some mental states M, and some cognitive or epistemic relation Φ, necessarily, if one is in M, then one Φs that one is in M.[5]

Historically influent instances of Φ, for example, are *knowing* or *justifiably believing*—the idea thereby expressed is usually called *self-intimation*. The relation of *being in a position to safely believe* is at least as interesting in this context. According to the target

of the reshaped argument, our beliefs about our current phenomenal states are waterproofed against epistemic luck, at least if they are based on careful introspection. In contrast to perceptual beliefs about the external world, introspective beliefs about one's current phenomenal states, so the idea goes, are especially secure in that, if they are true, they could not easily have been false. In my view, this is a veritable privileged access thesis, and, more importantly, an argument questioning that thesis is worth considering.

Therefore, let us revisit the case of Mr. Davis. Let t_0, t_1, ..., t_n be a series of times at one millisecond intervals from dawn to noon. Accordingly, call the situation in which Mr. Davis woke up "α_0" and the situation at noon "α_n." It is plausible to assume that somewhere in between there is a phenomenal watershed, so to speak: there are two bordering cases, α_i and α_{i+1}, such that in α_i, Mr. Davis feels cold and in α_{i+1}, he no longer feels cold.[6] Moreover, suppose for the sake of argument, that the content of Mr. Davis's beliefs is perfectly aligned with his thermal state such that, in α_i, Mr. Davis truly believes that he feels cold, whereas, in α_{i+1}, he no longer believes that he feels cold. Now, the crucial question is this: Does Mr. Davis, in α_i, *safely* believe that he feels cold?

At first glance, the right answer seems to be "no." The reason is provided by the following conditional that I assume will be acceptable to all safety theorists regardless of whether they adhere to a "zero tolerance" principle concerning false beliefs in any close worlds (like Williamson 2000 or Sosa 1999) or not (like Pritchard 2005):

A true belief b, actually held by a subject S, is unsafe if there is an extremely similar possible situation in which S continues to form b in the same way as in the actual situation, even though b is false.

In the case at hand, there certainly *is* such a situation: though, in actuality, Mr. Davis, at t_{i+1}, no longer believes that he feels cold, *he might well have continued to believe that he felt cold at t_{i+1}*. Recall that, similar to all of us human beings, Mr. Davis is not able to distinguish between his thermal experience at some point in time and his thermal experience one millisecond later. Thus, Mr. Davis is not able to distinguish between his thermal experience at t_i and his thermal experience at t_{i+1}. Therefore, it is eminently possible that he might have mistaken his thermal experience at t_{i+1} for his thermal experience at t_i, such that at t_{i+1}, he incorrectly believed that he felt cold. Let us call this latter situation "β_{i+1}." Certainly, β_{i+1} is not actual. However, it seems fairly possible. Moreover, β_{i+1} is extremely similar to α_i. From the perspective of Mr. Davis, for example, there is no difference between α_i and β_{i+1}: both situations feel the same for him. In β_{i+1}, Mr. Davis forms a belief that he feels cold in exactly the same way as he does in α_i, namely, based on his current introspective evidence. Thus, there seems to be no denying that in α_i, Mr. Davis *unsafely* believes that he feels cold.[7]

It appears then that we have before us quite a convincing argument against the claim that it is impossible that one fails to safely believe that one currently feels cold, given that (i) one currently feels cold, (ii) one attends to one's current feeling of cold as assiduously as possible, and (iii) one believes that one currently feels cold on the basis of one's introspective evidence. Note that this version of Williamson's argument copes without any margin for error principle—it merely makes use of a relatively

uncontroversial conception of safety. Thus, it cannot be rebutted by denying that safety is necessary for knowledge. It cannot also be dismissed as a sorites.

3 Safety regained through direct phenomenal concepts

However, there is a way to salvage the claim that Mr. Davis, in α_i, safely believes that he feels cold. Taking one's cue from Weatherson (2004), Berker (2008), and Ramachandran (2009), one might say that, in general, there is a "constitutive" connection between our phenomenal states and our beliefs about them that underwrites the claim that if Mr. Davis has done everything he can to decide whether he feels cold, then he believes that he feels cold only if he feels cold.[8] If the latter claim is true, there cannot be a situation such as β_{i+1}, that is, a situation in which Mr. Davis believes that he feels cold, though he actually does not feel cold.

There are several ways of unpacking what the constitutive connection between our phenomenal states and our beliefs about them might amount to. Weatherson, for example, claims that, when Mr. Davis "is in some phenomenal state, the very same brain states constitute both the phenomena and a belief about the phenomena" (2004, 379). In my view, Weatherson's claim is unfortunate in dialectical respects—for it presupposes materialism concerning the mind. Since materialism is a controversial doctrine, I think that the constitutivist fares better if he or she adopts a strategy that is based on the idea of *direct phenomenal concepts* in Chalmers' (2003) sense.[9] The constitutivist might then say that when Mr. Davis forms a belief as to whether he feels

cold, he uses a direct phenomenal concept of his present thermal condition, that is, a concept that somehow incorporates the thermal experience he currently undergoes.

To better understand how this strategy works, it may be instructive to regard direct phenomenal concepts as analogous to expressions that are formed with the help of quotation marks.[10] For example, Katalin Balog (2012) claims that there is a concept-forming mechanism, called "mental quotation," that operates on an experience and turns it into a direct phenomenal concept that refers to the type of phenomenal experience that the token exemplifies. Balog even uses a characteristic symbol for mental quotes, namely "*," to emphasize the analogy between linguistic and mental quotation. The only difference, she says, between linguistic and mental quotation "is that, unlike linguistic quotation, what is between mental quotes [...] is not a mental word but [...] an experience."[11] Thus, in light of Balog's account, there is a clear sense in which a conscious experience is *inside* a concept: just as the quoting expression " 'tree' " literally contains the word "tree," a direct phenomenal concept, say, of the feeling of cold literally contains that feeling.[12]

Let us now apply these ideas to the case of Mr. Davis. At each point in time, Mr. Davis holds some belief about his current thermal experience on the basis of his introspective findings. Recall that, in α_i, Mr. Davis feels cold, whereas, in α_{i+1}, Mr. Davis no longer feels cold, but presumably, is in an "in-between" state such that he neither feels cold nor feels hot. Let us represent the "cold" thermal experience that Mr. Davis undergoes in α_i as "↓" and the "in-between" thermal experience that Mr. Davis undergoes in α_{i+1} as "→." In α_i, then, Mr. Davis holds a belief that might be rendered as "I currently

undergo $*\downarrow*$," where "$*\downarrow*$" stands for a direct phenomenal concept that incorporates the very thermal experience that Mr. Davis undergoes in α_i. In α_{i+1}, however, Mr. Davis holds a belief that might be rendered as "I currently undergo $*\rightarrow*$," where "$*\rightarrow*$" stands for a direct phenomenal concept that incorporates the very thermal experience that Mr. Davis undergoes in α_{i+1}. Therefore, in both α_i and α_{i+1}, Mr. Davis holds a true belief about his current thermal experience. Moreover, at any point in time, Mr. Davis's belief is *safe*. This is because, for any actual situation α_x, there is no extremely similar possible situation in which: (1) Mr. Davis continues to form the respective belief in the same way as in the actual situation; and (2) the belief in question is false.

One might rub one's eyes in disbelief at this point: "Didn't you say that Mr. Davis is not able to distinguish between his thermal experience at some point in time and his thermal experience one millisecond later? So it *must* be possible that he mistakes the thermal experience that he undergoes in α_{i+1} for the thermal experience he undergoes in α_i!" However, the opponent here has not reckoned with direct phenomenal concepts. Consider what the idea of mistaking the thermal experience in α_{i+1} for the thermal experience in α_i could mean in terms of beliefs that involve direct phenomenal concepts. One might suggest that this idea amounts to a situation in which, at t_{i+1}, Mr. Davis forms a belief that might be rendered as: "I currently undergo $*\downarrow*$." However, recall that at t_{i+1}, Mr. Davis no longer feels cold: he no longer undergoes $\downarrow$, and instead undergoes $\rightarrow$. Thus, at t_{i+1}, it is not possible for Mr. Davis to form a direct phenomenal concept that incorporates $\downarrow$. Consequently, it is *impossible* that at t_{i+1}, Mr. Davis holds a belief that might be rendered as: "I currently undergo $*\downarrow*$."

One might condense these findings into the following conclusion: Necessarily, if S holds the belief "I currently undergo σ" at time t (where "σ" stands for some direct phenomenal concept that refers to some experience, namely, that very experience which is part of "σ"), then at t, S undergoes the very experience to which "σ" refers. This not only implies that one cannot hold a belief of the "I currently undergo σ" type if one does not undergo the experience that is part of the direct phenomenal concept one thereby uses, but also that beliefs of the "I currently undergo σ" type are *infallible*, that is, that there is just no possible situation in which someone holds a *false* belief of the "I currently undergo σ" type. From this, it follows quite trivially that there is no possible situation extremely similar to α_i (in which, recall, Mr. Davis holds the belief "I currently undergo *↓*") such that: (1) Mr. Davis continues to form the belief "I currently undergo *↓*" in the same way as he does in α_i, and (2) the belief "I currently undergo *↓*" is false. Therefore, Mr. Davis's belief about his own current thermal experience in α_i is undoubtedly safe.[13]

4 Confidence safety

However, there seems to be a way to avoid this conclusion: by slightly modifying the conception of safety. Instead of saying that *false belief* is sufficient for lack of safety, one might claim that *high confidence in a falsehood* is sufficient for lack of safety:

A true belief that p, actually held by subject S with degree of confidence c, is unsafe if there is an extremely similar possible situation in which S continues to have an at-most-slightly-lower degree of confidence $c*$ in the proposition that p, though p is false.[14]

In order to see the point of this modification, note that one can have a high degree of confidence in a proposition without believing it to be true. Let us assume, then, that the confidence threshold of outright belief is 0.8: if the degree of confidence in a proposition p is 0.8 or greater, then we have an outright belief in p. If, in contrast, the degree of confidence in p is less than 0.8, then belief in p has vanished.

Let us revisit, then, the example of Mr. Davis. It is obvious that in α_i, Mr. Davis has a degree of confidence in the proposition that is expressed by "I currently undergo $*\downarrow*$" that is sufficient for outright belief. However, since Mr. Davis's belief in this proposition vanishes within the next millisecond, it is natural to assume that the respective degree of confidence is only barely above the threshold for outright belief. Let us say that it is exactly 0.8. The crucial question is: how do matters stand in α_{i+1}? Since, in α_{i+1}, Mr. Davis no longer holds the belief "I currently undergo $*\downarrow*$," (instead, recall, he holds the belief "I currently undergo $*\rightarrow*$") his degree of confidence in the proposition expressed by "I currently undergo $*\downarrow*$" must be less than 0.8. However, since Mr. Davis held the belief "I currently undergo $*\downarrow*$" only a millisecond before, and, in addition, is not able to tell $\rightarrow$ from $\downarrow$, it seems reasonable to suspect that his degree of confidence in the proposition expressed by "I currently undergo $*\downarrow*$" is only *very slightly* lower than 0.8—say, 0,7999. Thus, it turns out that, in α_i, Mr. Davis holds a belief that is *not* safe; for there is an extremely similar possible situation, namely α_{i+1}, in which he continues to have an at-most-slightly-lower degree of confidence in the

11

proposition expressed by "I currently undergo $*\downarrow*$," although this proposition is false. It seems, then, that "even a constitutive connection between feeling cold and believing one feels cold is insufficient to vindicate luminosity" (Srinivasan 2015, 310).

The standard luminist response to this challenge is to doubt the modified conception of safety in terms of degrees of confidence. For example, Berker (2008) has claimed that the modified safety condition "deems as unreliable belief-forming mechanisms that appear to be as reliable as they could possibly be" (12). Regardless of the merits of this strategy, I think that the luminist can provide a more straightforward response. To understand this, let us take a closer look at the proposition central to our consideration: the proposition about his thermal experience that Mr. Davis believes in α_i. In the preceding paragraph, I have referred to this proposition as "the proposition that is expressed by 'I currently undergo $*\downarrow*$'." However, this is not quite accurate, because different tokens of "I currently undergo $*\downarrow*$" can express *different* propositions. Thus, it would be more appropriate to specify the proposition in question as "the proposition that is expressed by the sentence 'I currently undergo $*\downarrow*$' *as it is entertained in thought by Mr. Davis at t_i*." Let us call this proposition "p_i." One might say that p_i consists of Mr. Davis, time t_i, the thermal experience that Mr. Davis undergoes at t_i, and the relation of *undergoing*. Note that p_i does not change its truth-value across time. In particular, p_i is true in α_{i+1}. Thus, α_{i+1} does not qualify as a situation extremely similar to α_i (in which, recall, Mr. Davis has a degree of confidence in p_i that is sufficient for belief) such that: (1) Mr. Davis continues to have an at-most-slightly-lower degree of confidence in p_i; and (2) p_i is false.

Certainly, the proposition that is expressed by the sentence "I currently undergo $*\downarrow*$" *as it might be entertained in thought by Mr. Davis at t_{i+1}*—call it "p_{i+1}"—is false. However, mind you, this is *not* the proposition that Mr. Davis believes to be true at t_i. While the latter proposition, that is, p_i, consists, inter alia, of time t_i, the former proposition, that is, p_{i+1}, consists, inter alia, of time t_{i+1}. Thus, even on the modified version of safety in terms of degrees of confidence, Mr. Davis's belief about his thermal experience in α_i proves to be *safe*: there is no possible situation that is extremely similar to α_i in which the proposition Mr. Davis believes true in α_i is false and he nevertheless continues to have an at-most-slightly-lower degree of confidence in it.[15]

It must be noted here that the difference between p_i and p_{i+1} is extremely slight. Thus, in order to defend the claim that Mr. Davis's belief about his thermal experience in α_i is *not* safe, one might suggest a slightly modified version of the already modified conception of safety:

A true belief that p, actually held by subject S with degree of confidence c, is unsafe if there is an extremely similar possible situation in which S has an at-most-slightly-lower degree of confidence $c*$ in an at-most-slightly-different proposition $p*$, though $p*$ is false.[16]

However, even if we grant this once-more-modified conception of safety, it does not follow that Mr. Davis's belief about his thermal experience in α_i is unsafe. Recall that the thermal experience symbolized by "$\downarrow$" has ceased to exist at t_{i+1}. Because of this, it is no longer possible for Mr. Davis to form the direct phenomenal concept "$*\downarrow*$" at t_{i+1}. Accordingly, Mr. Davis is not capable of entertaining the sentence "I currently undergo $*\downarrow*$" in thought at t_{i+1}. It follows that it is impossible for Mr. Davis to grasp the

13

proposition p_{i+1} at t_{i+1}. Therefore, *Mr. Davis cannot have any degree of confidence in p_{i+1} either*. The proposition p_{i+1}, as it were, is just not on Mr. Davis's mental radar in α_{i+1}. Summarizing, α_{i+1} does not count as a situation extremely similar to α_i in which Mr. Davis continues to have an at-most-slightly-lower degree of confidence in the merely slightly different, but nonetheless false, proposition p_{i+1}.

One might object that it may be right that Mr. Davis cannot have any degree of confidence in p_{i+1}; however, this does not imply that there is *no proposition whatsoever* that might work instead. Consider, for example, the proposition that is expressed by the sentence "I currently undergo *that*" as might be entertained in thought by Mr. Davis at t_{i+1}—where the demonstrative refers to the "cold" thermal experience Mr. Davis underwent only millisecond ago, at t_i.[17] This proposition, call it "p^{that}," is clearly false at t_{i+1}. Nevertheless, it is plausible to assume that Mr. Davis has a degree of confidence in p^{that} at t_{i+1} that is at most slightly lower than the degree of confidence he has in p_i at t_i. Therefore, it seems that we have a situation that is extremely similar to α_i in which Mr. Davis continues to have an at-most-slightly-lower degree of confidence in a false proposition that is merely slightly different from p_i.[18]

In my opinion, this objection is based on an excessively liberal view about what should count as a *slight difference* between propositions. Note that in contrast to the difference between p_{i+1} and p_i, the difference between p^{that} and p_i not only concerns the time parameter, but also extends to the *mode of presentation under which the "cold" thermal experience in question is given to Mr. Davis*. I admit that it is difficult to explain the difference between the mode of presentation expressed by a direct phenomenal concept such as "*↓**" and the mode of presentation expressed by a demonstrative concept such

as "that." However, I think that it is possible to acquire at least an approximate idea by comparing the semantic behavior of these concepts across possible worlds.

Let us begin with the demonstrative concept "that" (as used by Mr. Davis at t_{i+1} to refer to the thermal experience he underwent only a millisecond before). Similar to all demonstrative concepts, "that" is a *rigid designator*, i.e. it rigidly refers to the item, whatever it is, that is the target of the corresponding demonstration. In addition to being rigid, "that" also exhibits some amount of situation-sensitivity: its reference might shift depending on the situation in which the speaker or thinker finds himself/herself. In a case where the situation in which Mr. Davis uses "that" is significantly different from the actual situation, "that" would rigidly refer to *another* experience. Consider, for example, a situation in which a slight tickle in his right big toe is the target of Mr. Davis's demonstration. If this situation were actual, "that" would rigidly refer to the slight tickle in Mr. Davis's right big toe.

Similar things can be said about standing phenomenal concepts that pick out the "cold" thermal experience Mr. Davis undergoes at t_i. Consider, for example, "feeling of cold." Plausibly, this concept might be paraphrased roughly as "the feeling that is typically caused in me in environments where temperature is low." Now, consider a situation in which human thermal experiences are systematically inverted, that is, the feelings we call "feelings of cold" become typically caused in environments where temperature is high, whereas the feelings we call "feelings of warmness" become typically caused in environments where temperature is low. If this were the actual situation, "feeling of cold" would rigidly refer to what we call "feelings of warmness."

15

Now, as Chalmers plausibly argues, things are different with direct phenomenal concepts.[19] They, too, are rigid designators: they rigidly refer to the experience they are composed of. However, direct phenomenal concepts are not situation-sensitive at all. The reference of "$*\downarrow*$," for example, does *not* shift depending on the situation in which Mr. Davis finds himself. Consider again the situation in which a slight tickle in his right big toe is the target of Mr. Davis's demonstration or the situation in which Mr. Davis has "inverted" thermal experiences. Even if these situations were actual, the reference of "$*\downarrow*$" would *not* change. In general, there is no possible situation in which the reference of "$*\downarrow*$" would change. In any case, "$*\downarrow*$" would continue to rigidly refer to the thermal experience of which it is composed. Chalmers sums things up by saying that direct phenomenal concepts are *super rigid*: they are not only rigid concerning their *secondary* or *subjunctive intension*, but are also rigid concerning their *primary* or *epistemic intension*.[20]

Now, given this difference in the modes of presentations associated with "that" and "$*\downarrow*$," the propositions p_i and p^{that} are *not* sufficiently similar. Certainly, they may be said to be similar in that they attribute the same property (*to be currently experienced by Mr. Davis*) to the same item (a particular feeling of cold). However, they are strikingly dissimilar in that they present this item through different modes of presentations. Thus, p^{that} bears a relation to p_i that is similar to the relation that the proposition expressed by the sentence "The morning star is shining" bears to the proposition expressed by the sentence "The evening star is shining." The bottom line is that p^{that} (as well as any proposition obtained by replacing "$*\downarrow*$" by a standing phenomenal concept) is not *slightly* but *considerably* different from p_i.[21]

16

5 A Pyrrhic victory?

In this section, I wish to address an objection that might and will no doubt be raised.[22] Recall the more general philosophical motivation behind Williamson's anti-luminosity argument: its purpose is to discredit the view that there are some states of affairs—or "conditions," as Williamson prefers to put it—such that the mere obtaining of them is sufficient for being in a position to know (or at least safely believe) that they obtain.[23] According to the view in question, these conditions form a so-called "cognitive home," that is, a point from which thought starts to develop its conception of the rest of the world.[24] Now, the objection is that my argument may have saved a cognitive home, but only at the cost of making it so small that it cannot be used as the thought's starting point.[25]

In order to make this objection more vivid, compare the beliefs "I currently undergo *↓*" and "I currently undergo a feeling of cold" as they might be held by Mr. Davis in α_i. Both beliefs are true, but only the first one is both safe and infallible. However, it seems that this is only an empty compliment—for, in contrast to the second belief, the first belief cannot be used as a starting point for any inference that goes beyond such trivialities as "I currently undergo an experience of some kind or other," or so it seems. The reason is that using an item as a representation of the type to which it belongs does not necessarily provide one with any clue concerning *which type it is* to which the item belongs. In this respect, the case of "I currently undergo *↓*" seems to be similar to the case in which someone lays his hand on top of his head and says "I am *this* tall."[26] As

one cannot infer from "I am *this* tall" how tall one actually is, one cannot infer from "I currently undergo *↓*" the type of experience that one undergoes. Moreover, it seems that "I currently undergo *↓*" cannot even rationally support (or confer justification to) any substantial belief about the phenomenal type of one's current experience. It appears, then, that beliefs such as "I currently undergo *↓*" may be safe and infallible whenever they are held, but only at the cost of being devoid of any epistemic value.

I must confess that, at times, I found this objection quite compelling. Nevertheless, I think that, in the end, it does not go through. The reason is that there is a crucial disanalogy between demonstratives and direct phenomenal concepts: while demonstratives can be successfully used to refer to something in the absence of any idea about what this something is like, any successful use of direct phenomenal concepts presupposes at least *some* idea about what the item to which it refers to is like. Consider a blindfolded subject who lacks any information about the environment she is in, but nevertheless randomly points with her index finger in some direction and says "*This* is a tree." Although the subject does not know to which object she points and thus lacks any idea about what this object is like, she successfully refers to the object, whatever it is, which lies at the end of the imaginary line that emanates from the tip of her index finger. In the case of direct phenomenal concepts, things are different. The reason is that one cannot even *form* a direct phenomenal concept unless one *has* the respective experience. Having an experience, in turn, means that *there is something it is like* for the subject to have this experience. Now, there being something it is like for the subject to have an experience implies that the subject has at least *some* idea of what it is like to have the experience in question—even if this idea is ineffable. Thus, one cannot form

the belief "I currently undergo *↓*" *unless one has at least some idea of what it is like to have a feeling of cold*. If this is correct, then it is simply not true that believing "I currently undergo *↓*" does not provide one with any clue concerning the phenomenal type of one's experience. Therefore, beliefs such as "I currently undergo *↓*" are *not* worthless in epistemic respects.[27]

I hope that the previous considerations also mitigate a related concern that might be raised. Recall that Williamson does not claim that there are *no* luminous conditions. He only claims that there are no *non-trivial* conditions.[28] As an example of a belief about a trivial condition in Williamson's sense, consider the case of Rudolf Lingens, an amnesiac lost in the Main Library at Stanford on July 4, 1970 afternoon. Even though Lingens does not know who he is, where he is, and what time it is, it is very easy for him to know what he would express by uttering the sentence "I am here now." The condition thereby expressed lies right before his eyes, so to speak. Lingens just cannot be ignorant or wrong about that. Thus, Williamson would concede that the condition Lingens believes to be true is luminous. However, Williamson would also insist that this condition is trivial, because the truth of what Lingens believes is compatible with a vast variety of possible situations—not just with the situation in which he actually is, that is, *being in the Main Library at Stanford on July 4, 1970 afternoon*, but also with *being in King Street, Boston, on the morning of March 5, 1770* or *being on Mars on December 12, 3040 at noon*.[29] The objection, then, is that Mr. Davis's belief "I currently undergo *↓*" is in the same boat as Lingens' belief—for it may seem as if the truth of what Mr. Davis's believes is compatible with a vast variety of possible situations as well.[30]

19

However, it is easy to see why this objection does not succeed. The reason is simply that the variety of situations compatible with the truth of Mr. Davis's belief "I currently undergo *↓*" is actually very, very small: It contains only situations in which the subject has a feeling of cold.[31] This is a consequence of the fact that "*↓*" is epistemically rigid. The epistemic rigidity of "*↓*," in turn, is a consequence of the fact that Mr. Davis—by undergoing the experience from which "*↓*" is built—directly grasps what it is like to feel cold. Now, a belief that is compatible with only a very small variety of alternative situations can hardly count as a belief about a trivial condition. Thus, the charge of triviality is unfounded.

6 Summary

This paper aimed to disprove Williamson's anti-luminosity argument, even if one grants the anti-luminist a refined conception of safety in terms of degrees of confidence. In order to make Williamson's argument as strong as possible, I first reshaped it in a way that avoids reference to the notorious margin for error principle—for reference to this principle considerably reduces the *prima facie* plausibility of the argument. Then, I outlined the discussion to the point where the constitutivist enters the stage. I argued that the constitutivist is well-advised to adopt the idea of direct phenomenal concepts, that is, phenomenal concepts that literally contain the experiences to which they refer. The natural response of the anti-luminist was to retreat to a conception of safety to the effect that high confidence in a falsehood is already sufficient for lack of safety.

20

However, it emerged that constitutivism backed by the idea of direct phenomenal concepts is immune against that strategy.[32] Finally, I discussed and rebutted two objections to the effect that I may have saved luminosity but at the price of rendering it uninteresting in philosophical respects.[33]

Institute of Philosophy

Goethe-University Frankfurt, Germany

References

Balog, K. (2012). "Acquaintance and the Mind-Body Problem," in S. Gozzano and C. Hill (eds.), *New Perspectives on Type Identity. The Mental and the Physical*. Cambridge: Cambridge University Press, pp. 16–42.

Berker, S. (2008). "Luminosity Regained," *Philosopher's Imprint* 8, pp. 1–22.

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

Chalmers, D. (1996). *The Conscious Mind. In Search of a Fundamental Theory*. Oxford: Oxford University Press.

Chalmers, D. (2003). "The Content and Epistemology of Phenomenal Belief," in Q. Smith and A. Jokic (eds.), *Consciousness: New Philosophical Perspectives*. Oxford: Clarendon, pp. 220–272.

Cohen, S. (2010). "Luminosity, Reliability, and the Sorites," *Philosophy and Phenomenological Research* 81, pp. 718–730.

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

Conee, E. (2005). "The Comforts of Home," *Philosophy and Phenomenological Research* 70, pp. 444–451.

Cresswell, M. J. (1985). *Structured Meanings*. Cambridge, MA: MIT Press.

Gertler, B. (2001). "Introspecting Phenomenal States," *Philosophy and Phenomenological Research* 63, pp. 305–328.

Gertler, B. (2012). "Renewed Acquaintance," in D. Smithies and D. Stoljar (eds.), *Introspection and Consciousness*. Oxford: Oxford University Press, pp. 93–127.

Horgan, T., and Kriegel, U. (2007). "Phenomenal Epistemology: What Is Consciousness That We May Know It So Well?" *Philosophical Issues* 17, pp. 123–144.

Hossack, K. (2002). "Self-Knowledge and Consciousness," *Proceedings of the Aristotelian Society* 102, pp. 163–181.

Hossack, K. (2007). *The Metaphysics of Knowledge*. Oxford: Oxford University Press.

Lehrer, K. (1997). *Self-Trust*. Oxford: Oxford University Press.

Leitgeb, H. (2002). "Review of 'Timothy Williamson, Knowledge and Its Limits'," *Grazer Philosophische Studien* 65, pp. 207–217.

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

Papineau, D. (2002). *Thinking about Consciousness*. Oxford: Oxford University Press.

Pritchard, D. (2005). *Epistemic Luck*. Oxford: Oxford University Press.

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

Smithies, D. (2012). "A Simple Theory of Introspection," in D. Smithies and D. Stoljar (eds.), *Introspection and Consciousness*, Oxford: Oxford University Press, pp. 259–293.

Sosa, E. (1999). "How to Defeat Opposition to Moore," *Philosophical Perspectives* 13, pp. 141–154.

Srinivasan, A (2015). "Are We Luminous?" *Philosophy and Phenomenological Research* 90, pp. 294–319.

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

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

Williamson, T. (1996). "Cognitive Homelessness," *The Journal of Philosophy* 93, pp. 554–573.

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

Wittgenstein, L. (1953). *Philosophical Investigations*. Oxford: Blackwell.

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

---

[1] Certainly, this sketch of the discussion is somewhat ahistorical. For example, Williamson (2000, 100) already anticipates the objection that there might be a constitutive connection between experiences and beliefs. However, as far as I know, it was Weatherson (2004) who first explicitly articulated this objection. Moreover, the refined safety condition in terms of degrees of confidence already appears in Williamson (2000). Srinivasan (2015) vindicates the refined safety condition against criticism by Leitgeb (2002), Berker (2008), Ramachandran (2009), and Cohen (2010).

[2] For arguments along these lines, see Brueckner and Fiocco (2002), Neta and Rohrbaugh (2004), Comesaña (2005), and Conee (2005).

[3] Cf. Wong (2008), Vogel (2010).

[4] Although she does not explicitly reshape Williamson's argument along these lines, my reformulation owes much to Srinivasan (2015); see especially p. 302 and pp. 306–8.

[5] I adopt this formulation from Smithies (2012, 264), where a "schematic version of the accessibility thesis" is formulated. Note, however, that unlike my formulation, Smithies' accessibility thesis has the form of a biconditional that combines a self-intimation claim and an infallibility claim.

[6] Certainly, this is not to say that in α_{i+1}, Mr. Davis feels hot. Note that *not feeling cold* is compatible with *neither feeling cold nor feeling hot*. Nonetheless, the assumption that there is a clearly defined boundary between cases in which Mr. Davis feels cold and cases in which he no longer feels cold is controversial. As both Stefan Reining and one of the journal reviewers have pointed out to me, proponents of ontic and semantic vagueness might claim that in α_{i+1}, it is indeterminate whether Mr. Davis feels cold. Accordingly, it would not be true to say that Mr. Davis no longer feels cold in α_{i+1}—for Mr. Davis still feels cold in α_{i+1}, but indeterminately so. Therefore, the appeal to vagueness seems to be a promising strategy to refute Williamson's argument at the very outset. However, I do not wish to pursue this line of thought any further, because in this paper, I am interested in another objection to Williamson's argument, an objection that does without any appeal to vagueness. Thus, for the sake of argument, I will accept the assumption that there is a clearly defined boundary between cases in which Mr. Davis feels cold and cases in which he no longer feels cold.

[7] Cf. Srinivasan (2015, 302).

[8] Cf. Weatherson (2004, 379); Berker (2008, 9); Ramachandran (2009, 665). See also Srinivasan (2015, 308).

[9] Cf. Chalmers (2003), particularly sections 3.1, 3.2, and 4.1. Similar ideas about phenomenal concepts can be found in Gertler (2001, 2012) and Horgan & Kriegel (2007).

[10] This analogy was originally used by Lehrer (1997), 167–170, and Papineau (2002), 116–121. Although Chalmers (2003) does not use the analogy of quotation, some of his remarks draw a picture of direct phenomenal concepts that almost resemble the quotational account. On p. 243, for example, he suggests "that the basis for a direct phenomenal concept contains within it a 'slot' for an instantiated quality, such that the quality that fills the slot constitutes the content." It is natural to think of this slot as the space between mental quotes.

[11] Balog (2012, 35).

[12] One may have reservations about the very idea of direct phenomenal concepts. It may seem puzzling, for example, that a direct phenomenal concept has a very short lifespan: it only exists as long as the experience from which it is built exists. Thus, in contrast to garden-variety

concepts such as "triangular," a direct phenomenal concept cannot be used repeatedly. Rather, it can be used only *once*. Because of this, one may think that direct phenomenal concepts are not concepts at all. I acknowledge these concerns. However, due to space constraints, I cannot provide a full-blown vindication of direct phenomenal concepts here, but refer the reader to Chalmers (2003) and Gertler (2001, 2012). If one continues to be skeptical about the very idea of direct phenomenal concepts after reading these papers, one may take my line of argument in a conditional sense: *If* direct phenomenal concepts exist, then there are beliefs about one's current phenomenal states that prove to be immune to Williamson's anti-luminosity argument.

[13] Cf. Weatherson (2004) and Berker (2008), who arrive at a similar conclusion. However, there are some differences between their considerations and mine. As I have already noted, Weatherson, for example, makes the rather strong, and thus implausible, claim that the experience of feeling cold and the belief that one feels cold might be constituted by the same brain state. Berker, in turn, assumes that "If one has done everything one can to decide whether one feels cold, then one believes that one feels cold only if one feels cold" (Berker 2008, 9). In my view, Berker is not quite correct here, as one can certainly believe that one feels cold even though one does not feel cold. The reason is that one might well use what Chalmers (2003, 239) calls *standing phenomenal concepts* in the process of forming one's belief.

[14] Cf. Williamson (2000, 97); Srinivasan (2015, 309).

[15] One of the journal reviewers has pointed out that the proponent of anti-luminosity might be able to avoid this result by spreading out the two relevant cases, that is, α_i and α_{i+1}, over different possible worlds. Imagine, for example, that there is a non-actual world that differs from the actual world only in that the timespan during which Mr. Davis's thermal feelings change gradually from cold to warm is shifted *only a millisecond back*. In this world, call it "w," there is a case in which Mr. Davis has the same thermal experience as he actually has at t_{i+1}, namely, an "in-between" state of neither feeling cold nor feeling warm. Since the whole series of cases is shifted a millisecond back, however, this case does not occur at t_{i+1} in w, but at t_i. Call this possible case "δ_i." Now consider the proposition that would be expressed if Mr. Davis entertained the sentence "I currently undergo *↓*" in thought in α_i and compare it with the proposition that would be expressed if Mr. Davis entertained the same sentence in thought in δ_i. Since the times are the same and nothing else has changed, it seems that these propositions are identical. It seems to follow, then, that there is an extremely similar possible situation, namely δ_i, in which Mr. Davis has an at-most-slightly-lower degree of confidence in the proposition he considers true in α_i, though that proposition is false in δ_i—for in δ_i, Mr. Davis no longer feels cold. Hence, the belief that Mr. Davis holds in α_i is *not* safe. However, this argument does not succeed in the end for reasons that I provide immediately. To anticipate: in δ_i, Mr. Davis is not

able to grasp the proposition he considers true in α_i; therefore, he cannot have any degree of confidence in it at all.

[16] Cf. Williamson (2000, 102).

[17] One might object that the "cold" thermal experience Mr. Davis underwent at t_i no longer exists at t_{i+1}, such that Mr. Davis cannot demonstratively refer to it at t_{i+1}. However, I think that it is quite easy to demonstratively refer to events in the past through one's memory. Thus, I see no fundamental problem in the idea of demonstratively referring to immediately bygone experiences.

[18] I owe this objection to one of the journal reviewers.

[19] Cf. Chalmers (2003).

[20] Cf. Chalmers (2003, 233).

[21] I am assuming here that propositions have an inner structure that exactly mirrors the conceptual structure of the sentences used to express them. I agree with the observation of one of the journal reviewers that this assumption is not uncontroversial. However, I think that there are good arguments for assuming that propositions are structured in a way that mirrors the conceptual structure of sentences—particularly if one takes propositions to be the contents of beliefs. See, for example, Cresswell (1985).

[22] I owe this objection to the three journal reviewers.

[23] Strictly speaking, it is incorrect to frame the issue in terms of conditions *tout court*—for one and the same condition might be luminous under one guise and non-luminous under another. Thus, Williamson occasionally speaks of conditions *presented under this or that guise* (cf. Williamson 2000, 94 f.). However, for the sake of brevity, I follow Williamson in leaving the reference to guises tacit.

[24] Cf. Williamson (1996, 573) who describes the view he criticizes as follows: "[T]hought initially engages with conditions whose *esse* is their *percipi*; if it later finds its laborious way to conditions of greater depth, it must do so from the starting point of that cognitive home."

[25] One of the journal reviewers appropriately rephrased this objection as "we've saved the foundations but lost the world."

[26] Cf. Wittgenstein (1953, §279).

[27] I admit that it is not at all easy to positively state what exactly the epistemic value of beliefs such as "I currently undergo *↓*" is; in particular, exactly what inferences they license us to draw or what other beliefs they can be used to evidentially support. The least that can be said, I think, is that the beliefs in question serve as a kind of interface between the realm of the phenomenal and the realm of the conceptual: they somehow lift phenomenal items onto the

conceptual level and thereby ensure that our conception of the world is constrained by our experiences.

[28] Cf. Williamson (1996, 557; 2000, 107–109).

[29] An alternative way of phrasing this is that the sentence "I am here now" as it is entertained in thought by Lingens in the Main Library at Stanford on July 4, 1970 afternoon is contingent according to its secondary proposition, but necessary according to its primary proposition. See Chalmers (1996, 63f.).

[30] I owe this objection to another one of the journal reviewers, different from the one who raised the previous objection.

[31] Again, an alternative way to phrase this is that the sentence "I currently undergo $*\downarrow*$" as it is entertained in thought by Mr. Davis in α_i is contingent according to *both* its secondary and primary propositions.

[32] One of the journal reviewers has noted that my overall argument in this paper has some resemblance to Keith Hossack's (2002; 2007) reaction to Williamson. However, as far as I can see, the differences prevail. For example, Hossack defends what he calls the "Identity Thesis," according to which "any conscious state is identical with knowledge of its own occurrence" (Hossack 2002, 174). The Identity Thesis, in turn, conceptually entails that *being in a conscious state* is a luminous condition in Williamson's sense. In contrast, I am not committed to a thesis as strong as Hossack's Identity Thesis. I do not argue for the claim that, say, *S* has a feeling of cold iff *S* has knowledge of that feeling. Rather, I argue for the claim that having a feeling of cold is sufficient for being in a position to know (or, more cautiously, to safely believe) that one undergoes a feeling of cold, provided that one picks out this feeling by a direct phenomenal concept. This is quite compatible with someone having a feeling of cold but lacking knowledge of that feeling.