

4-13-2011

Illuminating Anti-Luminosity

Tudor Alexandru Seserman
University of Missouri-St. Louis

Follow this and additional works at: <http://irl.umsl.edu/thesis>

Recommended Citation

Seserman, Tudor Alexandru, "Illuminating Anti-Luminosity" (2011). *Theses*. 222.
<http://irl.umsl.edu/thesis/222>

This Thesis is brought to you for free and open access by the Graduate Works at IRL @ UMSL. It has been accepted for inclusion in Theses by an authorized administrator of IRL @ UMSL. For more information, please contact marvinh@umsl.edu.

Illuminating Anti-Luminosity

Tudor Seserman

B.A., Philosophy, Hamilton College, 2009

A Thesis Submitted to The Graduate School at the University of Missouri – St. Louis in
partial fulfillment of the requirements for the degree
Master of Arts in Philosophy

April 2011

Advisory Committee

Berit Brogaard, Ph.D
Chair

Gualtiero Piccinini, Ph.D

Eric Wiland, Ph.D

Abstract:

In this essay I argue that the anti-luminosity argument Timothy Williamson presents in *Knowledge and Action* runs into problems in trying to establish its conclusion. Anti-luminosity is the position that we sometimes do not know our current mental states, for example I can feel cold and not know that I do. Williamson's argument initially seems plausible; however, it relies on an inadequately supported premise. Williamson needs to assume that the process by which we come to know our current inner states is fallible, but in doing so begs the question. Without assuming a fallible inner sense, the anti-luminosity argument looks like just another sorites paradox. Without problematically assuming fallible introspectors the anti-luminosity argument cannot establish its conclusion. One may think that this assumption is justified, but I argue that it is not. There are at least two other positions on the table that avoid a problem Williamson's position must face.

In *Knowledge and its Limits* Timothy Williamson argues compellingly for “anti-luminosity.” Anti-luminosity is the position that we sometimes do not know our current mental states, for example I can feel cold and not know that I do.¹ Two of the three dominant positions on self-knowledge, viz. Constitutivism² and Expressivism,³ would deny that this “anti-luminosity” argument is sound. The third position, viz. the Perceptual Model,⁴ readily accepts the conclusion for reasons independent of Williamson’s argument. Furthermore, as we shall see, Perceptual Model intuitions are needed to run the anti-luminosity argument, rendering the anti-luminosity argument unnecessary for the purposes of establishing the conclusion.

Williamson argues against luminosity, a minimal version of the traditional Cartesian picture. The Cartesian picture attempts to capture the very intuitive ideas that self-knowledge is peculiar, and privileged.⁵ Knowledge is peculiar if the method by which we know some x, is a method others cannot use to come to know that x (though they can come to know x by means of another method). The method, called “introspection,” by which an individual generally comes to know that she feels cold is different from the method others use in coming to know that she feels cold. Our knowledge is privileged if we can speak with a certain kind of authority. I may know that

¹ This paper is only concerned with self-knowledge about current mental states that have a phenomenal feel, states like pain, the basic emotions, temperature, and any other state differentiated from other mental states by phenomenology. I will not address beliefs, desires, and knowledge as a distinct mental state.

² See Shoemaker (1994), Wright (1998) and Moran (2001). Moran talks about beliefs and refuses to talk about sensations.

³ See Wittgenstein (1958), Finkelstein (2003), Bar-On (2004) and (2009), and Tye (2009).

⁴ See Armstrong (1968), Tye (2002), Byrne (2005), and Carruthers (2009).

⁵ The term ‘self-knowledge’ can broadly refer to the kind of knowledge we gain about ourselves from the first-person. Knowledge of our current mental states is only a kind of self-knowledge. When I use ‘self-knowledge’ I will generally be referring to knowledge of our current mental states. Philosophers who talk about self-knowledge generally approach introspection from a more conceptual perspective. Others approach introspection from a cognitive science perspective. I am more concerned with the conceptual issues.

Timothy Williamson wrote *Knowledge and its Limits*, but Timothy Williamson is more of an authority on its authorship.

Mental states need not be peculiar or privileged to be luminous, but our mental states are luminous if appropriately peculiar or privileged. Williamson defines “luminosity” as follows:

(Luminosity) For every case⁶ α , if condition C obtains in α one is in a position to know that C obtains.⁷

Contemporary accounts of self-knowledge try to explain luminosity in terms of peculiarity and privilege. On the Perceptual Model, self-knowledge is peculiar but not privileged. For Constitutivists self-knowledge is privileged and somewhat peculiar. According to Expressivists self-knowledge is privileged and not peculiar. On both Constitutivist and Expressivist accounts our mental states will be luminous.

I will first present Williamson’s argument and show that its fate rests with that of the Perceptual Model. I will then try to convey the *prima facie* plausibility of Constitutivism and Expressivism. I will conclude by introducing a problem for the Perceptual Model, which the other two positions avoid.

1 The Anti-Luminosity Argument

Imagine waking up in the morning and feeling cold. You warm slowly, and you know that you will be warm by noon because you are always warm by the temperature at

⁶ By “case,” Williamson means a subject at a time (Williamson 2005 p.94).

⁷ Conditions may be luminous in some cases, but there is no condition that will always be luminous when it obtains. (Ibid. p.95) Though Williamson formulates luminosity in terms of “position to know” I will assume that there is no significant difference between knowing something and being in a position to know. If anything, we know less than we are in a position to know. An argument using knowledge might be applied to a position formulated in terms of “position to know.”

noon.⁸ The temperature rises slowly, increasing by one degree every set period of time. The duration of time does not matter as long as you have enough time to develop a reliably based belief about how you feel. Little Brittany, just having learned the words ‘hot’ and ‘cold,’ asks with each change in degree whether you feel cold. Each time you answer “yes,” “no,” or “I don’t know.” If at any time you do not know whether you are cold, then your mental states are not luminous. Surprisingly, Williamson argues that even if you claim to know whether you are cold in every case there will be some states in which you do not know, because in these states your reports are not “safe” and therefore do not count as knowledge. In order to understand Williamson’s argument it will be helpful to look at the central notion of “Safety.”⁹

In *Knowledge and its Limits* Williamson argues at length that knowledge must be safe from error. The core idea is that you only know if you are not likely to be mistaken. This idea is expressed as a necessary condition on knowledge called Safety:

(Safety) You know p only if there are no close possibilities that p is false and you believe p [using the same method].¹⁰

⁸ The temperature at noon can be something absurdly high at which any normally functioning human being will consider very warm.

⁹ In the formal version of the anti-luminosity argument Williamson appeals to a margin for error principle that is derived from Safety. See Berker (2008) for a discussion of issues concerning deriving the margin for error principle from Safety. I take Safety to be the more general principle and our discussion of Safety will apply *a fortiori* to the margin for error principle that Williamson uses in the argument. I find Safety to be much more intuitive than the margin for error principle.

¹⁰ Alternatively, you know p only if p is true in all close possibilities in which you believe p [using the same method]. I give what Berker considers the course grained definition of Safety. Fine grained Safety is described in terms of levels of confidence or credences. I am skeptical of understanding Williamson’s argument in the way that Berker does because it does not distinguish between personal and epistemic credence. It is important to note that Williamson is an objective Bayesian who thinks that there are objective epistemic credences in a Carnapian heaven, and knowledge entails 1 prior probability. For Williamson only your evidence (or knowledge) can increase or decrease your credence in a given possibility. Credences can only vary with knowledge (or evidence) acquired. If one has knowledge, however they may feel about it, they know it with the same certainty that they know trivial truths. People that attempt to run the argument in terms of credences seem to have a subjective Bayesian understanding of credence, where someone’s credence can vary independently of what one knows (therefore collapsing the distinction between personal and epistemic credences the objective Bayesian recognizes). For Williamson, personal confidence tells us nothing about knowledge. See Berker (2008) and Ramachandran (2009) for fine-grained interpretations of Safety. See Williamson’s

Despite some reservations, for the purposes of this paper I will assume that Safety is correct in the unqualified form given above and that it is a condition on knowledge.¹¹

Safety was in part proposed to deal with epistemic luck cases such as Alvin Goldman's Fake Barn County. In Fake Barn County, there is one barn and many barn façades. One pertinent question is: Can someone, who does not know she is in Fake Barn County, know whether the only real barn is a barn? If she were taking a drive through the county and pointed to the only barn saying, "Look at that barn!" does she *know* it is a barn? It looks like she is just lucky, she could easily have pointed to the barn façade next door. The environment is not conducive to knowledge; there are too many close possibilities for error.

Carrying our discussion over to the anti-luminosity argument: When you try to answer Brittany, it would seem that your knowledge is safe only when there are no nearby cases in which you would falsely believe that you are warm. If you know that you feel cold at Dawn then at the very least you must believe that you feel cold and feel cold at Dawn₋₁ and Dawn₊₁. But we do not seem able to make distinctions between feeling cold and feeling warm in similar cases. You will be increasingly less sure that you are cold as you go from being cold to warm until you start feeling confident that you are warm. Wavering confidence does not tell us anything about what we know (see footnote 10), but our confidence wavers because we are generally unable to finely differentiate cases when the environment is full of chances for error. Wavering confidence seems to indicate that we do not think the process by which we come to know our mental states is

"Evidential Probability" chapter in *Knowledge and its Limits* for his objective Bayesian account. See Howson (1997) for issues that arise with objective Bayesianism.

¹¹ See Neta and Rohrbaugh (2004) for a proposed counter example.

infallible. There will be some cases where you could easily make a mistake. If you were able to identify when you go from cold to warm it would seem that you are merely lucky in the same way you would be lucky in Fake Barn County.¹²

Barn County can be adapted to better resemble the anti-luminosity argument. Imagine that you are driving through “Barn County (in progress).” In this county you first encounter just a plot of land, then you encounter the foundations of a barn, then a barn that is more complete and so on. There are small progressive changes from one plot of land to the next. Even if “barn” is not vague,¹³ you would have a very hard time identifying the first barn. Because you are driving, you are not sure if all the beams are in place, or maybe one barn is a few nails short of complete. As you are driving along the road you do what you can to make sure, but you get to a plot of land where you will not be able to safely say that you are looking at a barn or a barn in progress. You will be safe for a range of plots and only then can you claim to know. In this county, some barns are luminous cases for you and some are not.¹⁴

More formally, we can run the anti-Luminosity argument this way:

(Dawn) – A time at which you know that you feel cold.

(Noon) – A time at which you know that you feel warm.

(Luminosity) – For every case α , if condition C obtains in α one is in a position to know that C obtains.

¹² I take it that one could also argue that it is lucky that some beliefs are safe. Imagine Fake Barn Country where all the counties besides one are full of Barn facades. According to Williamson, your knowledge about barns in one county is safe but in all others it is not. It seems lucky that you are in the only county in the country where your knowledge is safe.

¹³ In other words, you know the necessary and sufficient conditions for “barn.”

¹⁴ Presenting the argument in this way seems to put a lot of pressure on your ability to come to know if something is a barn or not and conversely on your ability to discriminate between mental states. This is not the standard way to present anti-luminosity, but I think it is justified by the way Williamson hopes to maneuver away from turning the argument into a sorites argument.

We can sharpen ‘feels cold’ by using a physiological condition to resolve borderline cases. Let us assume that the subject of the process has no access to the technology needed to determine whether the physiological conditions obtains, and so is not in a position to know whether it does. These stipulations in no way weaken the argument for [the margin for error principle]. The considerations about reliability remain as cogent as before, for **they were based on our limited powers of discrimination amongst our own sensations**, not on the vagueness of ‘feels cold’ [my emphasis] (Williamson 2000 pgs.103-4).

(Safety)– You know C only if there are no close possibilities that C is false and you believe C [using the same method].

1. Given Dawn: there is such a time that you know that you feel cold.
2. From Safety: you do not falsely believe some time after Dawn (Dawn_{+1}) that you are cold, as long as that time counts as a close possibility to Dawn.
3. From Luminosity, 1, and 2: at Dawn_{+1} you know that you are cold.¹⁵

You can now run the argument again replacing “Dawn” with Dawn_{+1} . Running the argument a sufficient number of times will lead to the contradiction, $\text{Dawn}_{+N} = \text{Noon}$. By running the argument you will know at Noon that you are cold, but *ex hypothesi* you know that you are warm. Williamson argues that we should give up on luminosity to avoid the contradiction.

As presented, our mental states are not luminous because we cannot finely differentiate between them. This is a hotly debated question in the self-knowledge literature and we will now turn to some of the prominent positions. I will next argue that Williamson assumes the Perceptual Model in running the anti-luminosity argument, but before doing so I am going to address issues that arise with more sophisticated versions of the anti-luminosity argument. Most readers can skip this section.

1.1 The Anti-Luminosity Argument and Confidence

I have previously said that confidence is tangential to knowledge, which is not entirely accurate. Williamson thinks that in order to have the kind of belief that is associated with knowledge, we must be sufficiently confident of that belief. Confidence

¹⁵ To get to 3 we need the suppressed premise that for mental states to be luminous the subject, given that they have a belief at Dawn_{+1} , must either know that they are warm or that they are cold. If at Dawn_{+1} you simply believe that you are cold, luminosity also fails because it requires you to be in a position to know and not merely believe. One can challenge this premise by arguing that there are temperature states between warm and cold. However, the argument can be run with any adjacent mental states. For example if one thinks that one goes from feeling cold to feeling fine, one only need replace “warm” with “fine.”

thus enters the picture as a refinement on the simple version of the argument provided above. I think this refinement only works if you are willing to buy a considerable part of Williamsonian epistemology.

Williamson draws a distinction between a belief with a high degree of subjective probability and an outright belief. Beliefs with a high degree of subjective probability are beliefs about which we have enough evidence to consider probably true. On the other hand, beliefs we hold outright are beliefs we treat as true. According to Williamson, the most intuitive difference between the two is that outright beliefs can be used as a premise in a practical syllogism and beliefs with a high degree probability cannot.

Outright belief is the kind of belief that we are to associate with knowledge. If we only believe something is probable, we are not confident enough in that belief for it to be appropriate to knowledge. We can only be accused of unreliability about beliefs we hold outright. If a belief that we consider very probable turns out to be false, we are not unreliable because we have not fully committed to its verity. When we hold a belief outright, we do not allow for the possibility that we are wrong. Williamson argues that we only associate this second kind of confidence with knowledge.

One can now run the anti-luminosity argument in two ways different from the one offered in the previous section. One can run the argument by substituting outright belief for knowledge and offering a version of Safety for outright beliefs. We may want a Safety condition for outright beliefs in order to have reliable outright beliefs. If we have unsafe outright beliefs we would be more likely to become unreliable. We want reliable outright beliefs because we use them as the basis for action. Safety for outright beliefs would look something like this: You outright believe B only if there are no close

possibilities that B is false and you believe B [using the same method]. If there is a close possibility in which you believe B and B is false, you do not safely believe B in the actual case. This version of the argument will run into the same problems as the argument presented in the previous section. My criticisms of the anti-luminosity in next section will apply to this formulation *mutatis mutandis* insofar as it is structurally identical to the argument run in terms of knowledge.

The other possible version of the anti-luminosity argument run in terms of confidence bears less of a resemblance to the previously discussed anti-luminosity argument. When you start at dawn you have an outright belief that that you are cold, at a certain point you will go from having an outright belief to having a highly probable belief that you are cold. Because highly probable beliefs are not appropriate to knowledge, it is no longer possible for you to know if you are cold.

I have very little to say about this form of the argument. Intuitively, I do not think that knowledge requires outright belief. It is too stringent of a requirement on knowledge, it seems that most agents act on beliefs of which they are not fully confident. I will offer a brief response to the anti-luminosity argument with a strong reading of outright belief. The strong reading of outright belief holds that properly formed outright beliefs are strongly entailed by the evidence one has for that belief: two people with the same evidence, if they have a properly formed outright beliefs based on that evidence, will have the same outright belief. As an objective Bayesian, Williamson thinks that there are objective likelihoods given one's evidence for propositions. If I have some evidence to

support proposition x, but not enough evidence for me to know that x is true, there is a specific likelihood I must assign to x being true or be an unreliable reasoner.¹⁶

In the anti-luminosity argument, Williamson supposes that our reliably formed outright beliefs correspond to changes in our available evidence.¹⁷ As we go from warm to cold the evidence available to the individual changes, which corresponds to changes in outright belief. So what happens when we go from outright belief to a belief with a high degree of subjective probability? There seem to be two viable options, either the individual does not know how to appropriately interpret the available evidence or the she does not have enough evidence to form an outright belief. If the introspector is unable to use the available evidence, we must either assume a poor introspector or that “cold” is vague. If the introspector has the necessary evidence to form the belief that they are cold, it is hard to imagine why a person would be mistaken about being cold. To argue that a person, given the necessary evidence, would form an incorrect belief is to assume that people are poor introspectors (and in turn assume the anti-luminosity conclusion).

If the evidence available to the introspector is not sufficient enough to produce an outright belief, luminosity is not impinged. If the evidence available to the introspector does not conclusively point to the introspector being cold, the introspector has not failed at introspecting. It would be improper for an introspector to form an outright belief given evidence that does not conclusively point to her feeling cold. Unless we assume that people are poor introspectors, there is nothing wrong with forming a highly probable belief if the evidence calls for it.¹⁸

¹⁶ See the “Evidential Probability” chapter in *Knowledge and its Limits* especially pages 211-3.

¹⁷ For instance see the bottom of page 99 in *Knowledge and its Limits*.

¹⁸ Williamson may consider a case where the conscious evidence we have is not sufficient for us to know whether we are cold, a case in which we are not in a position to know our mental state (we are missing some evidence that is not consciously available to us at the moment and we need in order to be able to know whether we are cold). We are not in

Williamson might argue that we have good reason to think people will form improper beliefs given their evidence: because people are generally poor reasoners, introspectors will be unable to identify the point at which it is proper to have an outright belief given the evidence. There is plenty of evidence to indicate that people's beliefs are often improperly informed by their evidence, people are at times irrational. Even though we can acquire evidence for these kinds of errors in judgment, I am skeptical that on the Perceptual Model we have good reason to think that people are unreliable in introspecting. I will turn to these worries in the last section of the paper.

2 The Perceptual Model

According to the Perceptual Model, we have an inner sense with which we acquire knowledge about ourselves. Traditionally, this inner sense was thought to be an infallible organ, but contemporary versions support a fallible sense. Let us briefly look at a simplified traditional Perceptual Model before we turn to contemporary models.

On a caricatured Cartesian model of the mind, our mind is like an attic full of atomistic mental entities. When we need to know whether we feel cold, we use our flashlight (symbolizing awareness) to rummage through the attic to find the box that contains that information. On this model, self-knowledge is peculiar because nobody else has a flashlight with which to rummage through our mental attic. Introspection is privileged because we are infallible about the content of our "temperature box." We are the only people that can take a look at the box and only we are infallible about its contents. If one operates with this robust picture of introspective prowess, Williamson

a position to know because we lack evidence that is not available to us. I think this route is mistaken. Whatever evidence we can have for feeling cold ought to be able to available to consciousness. We can make mistakes in understanding that evidence or not being aware of it, but we should no require other evidence that is not available to know whether we are cold. Insofar as being cold is a conscious mental state all evidence we can have ought to be available to consciousness.

may be refuted. The Cartesian can hold Safety and argue that self-knowledge is safe in the way Safety requires – for any given mental state there are no close possibilities for error. Because your beliefs are infallible, only beliefs about which you are infallible can count as close possibilities. In the anti-luminosity argument, if your beliefs at Dawn₊₁ and Dawn₋₁ were not infallible they would not count as close possibilities. The Cartesian can also reject Safety by arguing that we know our mental states in such a way that we are guaranteed to know even if the environment is teeming with possibilities for error.

With the rise of psychology, people have moved away from the Cartesian picture. Empirical studies show that we are generally unaware of what stimulates a response, our response to stimuli, and how a stimulus affects a response.¹⁹ Philosophers who adopt contemporary versions of the Perceptual Model give up Cartesian infallibilism. For them, introspection is still a perceptual process but we are no longer infallible in our self-reports.

Working through an example, pain, will bring out the major facets of this position. You have a dull pain in your left knee that persists for some time. On the contemporary account, you are in pain as long as your knee is in pain. At times you may be thinking about the weather and completely forget that your knee is in pain, but you are still in pain. If somebody were to ask you, “Are you in pain?” you would need to look inward and see if “knee pain” is one of your mental states. Sometimes when you look inward you are mistaken. Suppose that you were out running, and adrenaline masks a pain in your knee. When you look inward you do not see that you are in pain, and you mistakenly come to the conclusion that you are not in pain. On the contemporary

¹⁹ See Nisbett and Wilson (1977) for a classic treatment of this topic. It should be noted that the psychologists are generally interested in the kind of self-knowledge that does not interest us here.

Perceptual Model, introspection is just like any other modality you use to gain knowledge. It is peculiar only insofar as any of our senses are peculiar. You cannot use my vision to see a mountain goat anymore than I can introspect your pain. Introspection certainly is not privileged – doctors are often in a better position to know whether you are in pain.

If we treat introspection as simply another modality, anti-luminosity is very appealing. It is generally accepted that we often make mistakes when we see, hear, or taste. We would never be tempted to consider these other modalities infallible. If inner sense is another modality, we similarly have no reason to consider it infallible. Like the other modalities inner sense is fallible, but sometimes reliable. The Perceptual Modal has a lot of intuitive appeal if we treat introspection as another modality. We will see that Williamson must assume the Perceptual Model's fallible inner sense in order to run the anti-luminosity argument. He must assume that we are fallible when it comes to our sensations, but if he does so he gets anti-luminosity for free.

First, Williamson can assume that our inner sense is fallible and thus sometimes reliable.²⁰ Inner sense can be reliable in one of three ways: always, sometimes, or never. If, as the Cartesian argues, our inner sense is always reliable we would either be safe in all self-knowledge assertions or we would know our mental states even in inhospitable circumstances. Cartesians would deny that Safety plays a role in the argument. The anti-luminosity argument would not work if our inner sense were always reliable.

On the other hand, if our inner sense is never reliable, Williamson could not establish two premises in the argument: Dawn and Noon. In order to establish these two

²⁰ Running the argument with a margin for error principle instead of Safety seems more straightforwardly question-begging. We would only need a margin for error if we were unreliable in introspection. But if we assume we are unreliable in introspection, there is nothing left to prove.

premises, there must be a time at which we know that we are cold and another at which we know that we are warm. These two premises cannot be established without an inner sense that is reliable enough to generate knowledge in at least those two states.²¹

Obviously, if our inner sense is always wrong, Williamson would not need the anti-luminosity argument. Assuming that inner sense is sometimes reliable is the only option open to Williamson.

As we shall see, if Williamson does not assume anything about the reliability of introspection, the argument would simply become a sorites paradox.²² For the sake of argument, let us assume that Williamson can establish Dawn and Noon without assuming anything about the reliability of introspection. As we go from Dawn to Noon we are again to unable to identify the exact moment we stop being cold and start being warm. What accounts for this deficiency? The vagueness of ‘cold’ and ‘warm’ seems the most likely candidate, just as it would in a standard sorites paradox. I can finely distinguish between grains of sand, I know if I have two or ten grains of sand. In the same way, unless we assume a fallible introspector, we should be able to know our sensations. In the first case I do not know when I have a heap, and in the second I do not know when I start to feel warm.

Williamson argues that the problem persists when we sharpen, or make precise, the vague terms. Why think that? Sharpening the terms in standard sorites cases solves

²¹ I will return to this point in the last section of the paper.

²² For sorites paradoxes see Williamson (1994). Sorites paradoxes are paradoxes that arise when we consider vague terms such as ‘heap.’ One grain of sand is not a heap. Adding one grain to that grain will not make it a heap, but what if we add another and another and so on. How many grains of sand does it take to get to a heap? There seems to be no clear line between heap and not heap.

It would be bad if the anti-luminosity argument were merely a sorites paradox for several reasons:

First, Williamson denies that it is a sorites paradox: “The [anti-luminosity] argument is not just another sorites paradox” (Williamson 2009 pg.104).

Secondly, insofar as some mental states are non-vague, there will be some mental states that are luminous. Friends of luminosity can live with some vagueness.

Third, people generally are not driven to Williamson’s position in sorites cases.

the problem. For example, let us see what changes when we sharpen 'heap' so that any three grains of sand is a heap.²³ Insofar as people can count the grains of sand they ought to be able to know when they encounter a heap. A person may be deficient in counting grains of sand, but that is entirely a separate issue. If we sharpen 'cold,' the anti-luminosity argument would equally lose any force we may think it had. If I know what it feels like to feel cold, or warm, than I should be able to know whether I am cold, or warm. The problem only emerges if we assume something about our ability to distinguish between our mental states, thus Williamson is forced to assume something about our inner sense.

If we sharpen the relevant terms in the anti-luminosity argument and avoid assuming a fallible inner sense, the argument becomes less convincing. The argument would go something like this: As you go from Dawn to Noon at some point you will stop being cold and start being warm. Assuming there is a sharp cutoff, will you be able to know when you have reached that point? The question cannot be decided, without knowing something about our ability to distinguish between mental states. The question can only be settled by knowing something about the degree to which we can reliably introspect. Yet, if we need to know whether we can reliably introspect to run the anti-luminosity argument, we do not need the anti-luminosity argument to know anything about our ability to introspect.

Williamson can perhaps be interpreted in a different way, but as we shall see he will still need to assume the Perceptual Model in at least a limited form. The argument can be run by only assuming that people are poor judges of their sensations, while leaving

²³ The larger the number the harder it may be to identify a heap, but given enough time to form a reliable belief anyone should in principle be able to identify a heap.

open the status of non-sensational states. This question is not straightforwardly question-begging because Williamson is trying to establish a claim about knowledge based upon data about our mental lives. In other words, Williamson can without begging the question assume that people are poor at distinguishing between sensational states such as ‘feel cold,’ as long as we treat non-sensational states like knowledge as a different kind of mental state. The argument would thus need to show that unreliability in knowing sensational mental states should lead us to think that we are equally unreliable with non-sensational state. Insofar as this argument would need to assume that we are unreliable introspectors, it assumes the Perceptual Model in a limited form. If it turns out that we are reliable introspectors of our sensations, his argument would be unsound.

Williamson may avoid begging the question in another way. He clearly assumes that we are limited in our ability to discriminate between mental states, see the quote in footnote 14. He may argue that assuming something about our ability to discriminate between mental states is different from our ability to know those mental states. It does not seem like this distinction would avoid the question-begging charge, insofar as being able to distinguish a mental state from others seems to be a necessary condition for knowing a particular mental state.

The problematic nature of this distinction is brought out by the structure of the anti-luminosity argument. The argument asks a subject to report whether or not they are cold at particular times. It is asking about a whole body state, and not whether some limb is cold. It seems absurd to think that a person can have two whole body states at the same time. You cannot be both cold and warm at the same time (part of you may be cold when you are warm and vice versa). Because we cannot be in two whole body states at the

same time, we cannot directly compare what it feels like to be cold now with how we felt at another time. We cannot compare various ways of being cold at the same time by being aware of them simultaneously as we would something like color samples. We may, perhaps, compare our current mental state with the memory of another, but nobody thinks we are infallible about our memories. When we try to discriminate between our current mental state and some past mental state we do so over a period of time and compare the reports we give at the different times. Being able to discriminate between mental states in this way requires that we know what mental state we are in at a particular time. I first need to know the mental state that I am in at time t to be able to know if it is different from the mental state I am in at t' . If I report my mental state at t in a different way than I report my mental state at t' , then I report them as different. If we were to assume that I am unable to discriminate between mental states in this way, we would have to assume at either t or t' that I do not know what mental state I am in. Assuming that I do not know my mental states begs the question.

One could think that we would fail at distinguishing between mental states even if we know which states we were in at the time: I know my mental state at time t and t' but at t' , for whatever reason, I fail to know whether my mental state at t' is different from my mental state at t . This assumption, though unproblematic, could not play the appropriate role in the anti-luminosity argument. Williamson needs to assume something about our ability to discriminate between sensations in order to show knowledge of our sensations can at times be unsafe. My knowledge at t' can be safe even if fail to know whether it is the same mental state as the one at t . My knowledge would be safe insofar as if I were in t I would know what mental state I am in, I would not have false beliefs in

nearby cases. This is not the kind of discriminatory failure Williamson assumes in running the anti-luminosity argument because it cannot do the work he needs it to do.

The anti-luminosity argument can be run using mental states that can be compared side by side, for example using the color spectrum. A subject is presented with reddish cards and he is asked if they are the same shade. Let us suppose that our subject gives a false answer, thus he has failed to discriminate between the shades. Is there any sense in which he knows which colors he is seeing? I think the answer is obviously no. If he is wrong to think that the colors are the same than he does not know at least one of the colors has obtained. He has a false belief about at least one of the colors, namely the belief that leads him to conclude that the color he sees on one card is the same as the color that he sees on the other. If we understand failure to discriminate in this way, anti-luminosity follows straightaway.

Two thesis unite contemporary versions of what I have been calling the Perceptual Model: 1) we know our sensations through perceptual like process, and 2) this process is fallible. I have spent the majority of this section arguing that Williamson assumes the second of these theses. I will now briefly turn to how the first thesis factors into Williamson's argument. Though the argument is rife with Perceptual Model-type language, but he does not explicitly endorse such a view though he argues against Constitutivist and Expressivist-like views. He does talk about the evidence we have for sensation when he discusses beliefs on page 99 of *Knowledge and its Limits*.²⁴ It may be that he does not endorse the Perceptual Model's first premise, despite appearances, but Perceptual Model seems like best theory to justify Williamson's intuitions he wants.

²⁴ We can assume from this he does not hold a Theory-Theory view.

It might be that Williamson has good reason to assume the Perceptual Model, fully or in modified form. On the full Perceptual Model, we get anti-luminosity for free. On a limited version, the anti-luminosity argument may go through. We may have reasons to hold the anti-luminosity conclusion, but the anti-luminosity argument should not get us there. I think that the Perceptual Model is mistaken; therefore I want to introduce some rival accounts that ought to dissuade us from simply assuming the Perceptual Model. If these accounts have some *prima facie* plausibility, the anti-luminist has some work to do.

3 Constitutivism

Constitutivists, true to their name, think there is a constitutive connection between our mental states and our beliefs about them.²⁵ I am cold if and only if I believe I am cold. One can put the idea common to different Constitutivists this way: C obtains in α iff α is in a position to know that C obtains. Where C is a current phenomenal mental state.

The Constitutivist disagrees with the Perceptual Model; there are no unfelt pains. When you go running and your body produces adrenaline, you are no longer in pain. Pain is not masked by anything new – rather, pain no longer exists.²⁶ You have as many sensations as you know you are having. Obviously on this account you can have pains that the Perceptual Model would not recognize as pains. You may wake up one day and have your head hurt. You go to the doctor, and the doctor tells you that there is nothing wrong with you. The Constitutivist has a straightforward response, whereas the

²⁵ Berker makes a similar argument, but does not present Constitutivism as a plausible model for human self-knowledge.

²⁶ There is a division in the Constitutivism literature. Some find it plausible for beliefs and other more complex phenomena, but implausible for sensations. Others find it plausible especially in the case of sensations. In this paper, I am focused on sensations.

Perceptual Model has some work to do. The Constitutivist assumes that if self-knowledge is privileged, you are a better judge of your mental states than anyone else.

There are many ways to run the Constitutivist argument, with at least two decision points worth mentioning. As a Constitutivist, one can make either a metaphysical or semantic claim. For the metaphysical Constitutivist, a certain mental state is present when and only when you are in pain. For the semanticist, pain is defined as those states that a person (or group) is willing to consider pain.²⁷ I cannot be wrong when I say that I am in pain because being in pain is the same as ascribing pain to myself.

Also, one must say something about the constitutive connection. There seem to be at least three ways one can flesh out the constitutive connection between mental states and our expression of those mental states. Depending on how one comes down on these issues, self-knowledge may be peculiar. First, merely being aware of pain would be sufficient for being in pain, given that being aware of pain (in the appropriate way) is being in pain. On this account any sentient being could be in pain if they feel pain. Second, being aware of pain in a conceptually laden way is needed to form a constitutive connection. On this account, infants and animals, insofar as they lack concepts, would not feel pain in the same way that we do. Lastly, one could argue that one has to avow pain in order to be in pain (there is no private language). Constitutivism's plausibility might depend on how you come down on these issues, but I think the tools to scrape together a plausible account of the position are there. Let us turn to one last issue that may impinge upon Constitutivism's *prima facie* plausibility.

²⁷ By semanticist I have in mind some one with deflationary tendencies willing to say the meaning of a word is whatever meaning we give, a "meaning is use" kind of person.

On Williamson's account constitutive state will be uninteresting because they are trivial. There is a sense in which mental states on the Constitutivist's account are trivial, but they certainly are not uninteresting insofar as they can factor into practical judgments. On the Constitutivist picture mental states can be used as the basis for action. If I think that I am in pain I will go to the doctor, whereas if I am happy I try to maintain that state. A runner full of adrenalin does not need to go to the doctor to cure his pain. He needs his knee fixed, but requires nothing for the pain (even though on the Perceptual Model he might). The Constitutivist can maintain the interesting connection between mental states and action.

Unlike the Perceptual Model, Constitutivism posits a close relation between mental states and our knowledge of them. There are many ways to formulate this position, but on any formulation, mental states are luminous. We will now turn to another plausible alternative to the Perceptual Model.

4 Expressivism

For Expressivists, self-knowledge is immune to error because there is no room for error between being in a certain mental state and the expression of that mental state.²⁸ We sincerely avow that we are in a certain mental state only when we are in that mental state.²⁹ Avowals are akin to verbal behaviors. When in pain, we exhibit certain behaviors; when we are happy, others. Avowals express our current mental state, but we also use them to say something true about our current mental states. When I was younger I cried when I was in pain, now I say, "I am in pain" (or something less stilted and more

²⁸ Some Expressivists deny that the information we gain upon introspection is factive. Insofar as we are discussing self-knowledge I mean to only be discussing those Expressivists that take the information gained upon introspection to be factive.

²⁹ One can of course lie about being in a certain mental state, but here we are interested in self-knowledge and not what others know about our minds.

colorful). In both cases I both make an assertion and express the pain. My attribution does not make me feel pain, but I can offer a truthful avowal because I am in pain.

The Expressivist shares many of the Constitutivist's intuitions and Expressivism is best understood as Constitutivism inverted. Constitutivists think there is a very close connection between a mental state and knowledge of that mental state. However, they take the mental state to be primary. According to Expressivist we only have knowledge about our mental states because we have those mental states. However, according to Constitutivists we have mental states because we are in a position to know them. They put the emphasis on knowing the mental state. Expressivists, on the other hand, take the expression to be most important. We are able to express our sensations because we have those sensations. Expression is the public part of a sensation.

Expressivists think we can make mistakes when we report our current mental states, but mental states are still luminous. For instance, I may say that a car is blue when really it is red. I do not see the car as red; 'blue' is merely the word I utter when I see the car. My report is false because the car is red and not blue. I express what I know about myself poorly, but my mental states are still luminous to me. I have not failed in knowing my sensation just because I fail to know what is in the world. Even though I expressed a falsehood, it is not a failure of self-knowledge.³⁰ The Expressivist holds that only things I know about myself are privileged. They need not claim that I have the right words to express that knowledge.

If I am in the presence of a duck, I can identify it as a *that*, as a discrete object. I see the duck even if I do not have a word for it. The same is true even if I have the wrong word. I can call a duck "alligator," but there is still a sense in which I have correctly

³⁰ Tye (2009)'s discussion on ps. 188-90 is quite helpful in elucidating this point.

identified the duck. If I were to call the duck an alligator in all close possible situations, it would seem that there is nothing wrong with what I am seeing. It is not a deep mistake. If I said, "That is an alligator," it would be false only insofar I am not a competent speaker of the language. Nothing would be revealed about my ability to identify ducks.

Expressivism, like Constitutivism, posits a close connection between mental states and our knowledge of them. This close connection helps these positions avoid a problem the Perceptual Model does not as easily avoid. We now turn to that problem.

5 A Problem for the Perceptual Model

The Perceptual Model faces a problem that both Constitutivism and Expressivism avoid: How do we know if people have self-knowledge? How do I (or you) know you are telling the truth when you say that you are cold? Some kinds of self-knowledge claims are readily tested. You can test whether someone is as tall or smart as they think they are. You can probably also test attitudes and beliefs by the same methods. Behavior inconsistent with a sincere asserted dislike for vanilla will show the claim to be false. Even though we can test a great amount of self-knowledge claims, the kind of self-knowledge with which we have concerned ourselves in this paper is not testable by the same methods.

I will argue that on the Perceptual Model we cannot know if an individual's sensational self-knowledge claims are true without the use of luminous cases. Because the Constitutivist and Expressivist deny any gap between mental states and knowledge of those mental states, they will simply reject the possibility of mental states that are not open to public scrutiny. Insofar as the Constitutivist and Expressivist positions considered here posit that we are infallible about certain mental states, they ought to have something

to say about why we are infallible about certain mental states. The account given of infallible sensations will vary with the different positions. Roughly they will say that there is a closer connection between certain mental states and their public aspect than there is amongst other mental states. We are infallible when there is no gap between the state and its public aspect. The mental state and the public aspect do not come apart as they do on the Perceptual Model.

The most salient way to test a self-knowledge claim requires the use of a “baseline.”³¹ A baseline is an unambiguous case, against which the truth of other cases can be tested. In general the baseline is something public against which we can measure the claim we want to test. If we want to make sure a meter stick is a meter, we go to Paris and compare it to the original one. We cannot do the same with inner states because on the Perceptual Model they are peculiar; they are not public. In order to test inner states we must tie them to something public, generally a kind of behavior, which we can test. If the Perceptual Model only uses a public baseline it will be unable to explain a host of cases, but if it moves away from a public baseline it must embrace skepticism about the mental.

We can develop (and naturally have developed) a baseline for whether an individual prefers vanilla to chocolate. The baseline would be something like: if an individual is offered chocolate or vanilla, she will choose vanilla *ceteris paribus*. If a person claims to like vanilla but the majority of the time she chooses chocolate over

³¹ There may be others way to test the self-knowledge, but this is the most salient and the one most people indicate will work. This argument is derived from arguments Wittgenstein makes in the *Philosophical Investigations*.

vanilla we can say that she does not know herself. Such people are ignorant of their preferences.³²

How would we develop a baseline for something like “sees red?” We can perhaps say something like: a person sees red and not green when they can reliably distinguish between color samples: if presented with two red samples she would know that green is not present, if presented with red and green she would be able to consistently distinguish between the two. What can we conclude from this? We can certainly conclude that she is able to distinguish between two hues insofar as we are able to measure the public performances. If the Perceptual Model only uses public baselines to evaluate self-knowledge, it will have problems explaining cases where this kind of public behavior does not match the sensations one is having.

Inverted spectrum worries are well known. Locke famously gave birth to the idea that we could imagine a person whose color wheel is inverted.³³ For instance, if I see something as red then a person whose spectrum has been inverted will see it as another color. People with inverted spectra are functionally identical to their counterparts: they will give the same answers in similar circumstances; they will make the same common mistakes; they will react to colors in standard ways (e.g. they will say rooms painted white are larger than rooms painted in dark colors); and they will display the same microbehavior. The possibility of inverted spectrum cases presents a challenge to the Perceptual Model. If people have inverted spectra we cannot develop a baseline for “seeing red.” Insofar as we require a public baseline for judging introspective reports, we

³² I imagine such cases are very common. See Hurlburt and Schwitzgebel (2007) for a discussion and Hurlbert (1993) for an example of a man that professes not to be angry at his children, but records numerous times he is angry with his children when asked to record moments of his life.

³³ For this discussion I find it helpful to assume color realism, though I have no particular stance in that debate. My discussion will ultimately hinge on an area where realism is much less controversial.

could not distinguish between people that have different sensations while functionally identical to those without inverted spectra.

Inverted spectrum cases create problems for the Perceptual Model because they are cases in which there is a mismatch between the sensation one has and what is in the world. If we use a public baseline the self-reports of people with inverted spectra will come out true, but on the Perceptual Model these reports should come out false. According to the Perceptual Model, people with inverted spectra are able to distinguish between different colors, but they are not able to distinguish between those colors *qua* those colors. A person with an inverted spectrum does not see red as we see red thus she does not know the sensation she is having. She thinks that she has the sensation of red when she is looking at a red object, but really she has the sensation of green. If sensations are private, as they are on the Perceptual Model, and if sensations do not differ in their public manifestation, as they are in inverted spectrum cases, then the Perceptual Model is unable pick out an error it itself predicts. If we use the public baseline to evaluate people with inverted spectrums, their reports will come out true, but according to the Perceptual Model the reports should come out false because the introspector has failed to correctly report her sensation. Whereas inverted spectrum cases are controversial, the same point can be made by looking at less exotic cases.

The Perceptual Model should be able to explain other cases in which there is mismatch between the sensation and the world, such as when a subject is hallucinating. Antoine, having spent a fair amount of time in the desert without any water reports seeing a sheep. Amelia, his companion who has been hoarding water, does not see the sheep concluding that Antoine is hallucinating. She does not deny that Antoine is seeing a

sheep, but she does realize that there is no sheep for Antoine to see. Amelia does not question that Antoine is seeing a sheep; she finds no reason to question his self-report based on the fact that there are no sheep present. Why not? Insofar as Amelia considers Antoine a truthful individual, she does not deny his introspective report despite contravening public evidence.

Even though Amelia believed Antoine when she knew him to be a truthful person, should she believe him in this case if he is an affirmed liar? Obviously, she should not. Should she deny that he sees a sheep? Again the answer ought to be no. Even though we know Antoine is an affirmed liar the evidence that we have allows for the possibility that Antoine is not lying in this particular case. We know that people stranded in the desert without water sometimes see mirages. What if Antoine were in his house drinking water, would we then have enough reason to deny that he is seeing a mirage? Again, the answer should be no. The evidence we had in the original case, and have slowly stripped away, is evidence for whether Antoine is offering us a truthful introspective report. The evidence only influences whether Amelia should think that Antoine is presenting his condition truthfully, not evidence that goes to show whether Antoine is indeed having a hallucination.

One way to bring this last point out is to ask whether our evidence would convince Antoine to recant. Amelia can tell Antoine: that he is an affirmed liar, that he is not suffering from dehydration, and that he has not done anything that generally induces hallucination. Antoine should still say that he sees a sheep if he indeed sees a sheep.³⁴ Antoine should not have the same worries about whether he is truthfully offering his

³⁴ Cases of self-delusion present a problem only if the image of sheep is not part of the self-delusion. I take it that if Antoine is deluding himself into seeing a sheep, part of that delusion will be imaging a sheep.

report. A public baseline can only get us so far, we are really interested in whether someone is having a sensation when they say they are. To determine this, we need a baseline which tracks whether someone is having a certain sensation when they say that they do. By giving up only evaluating introspective reports with a public baseline the Perceptual Model can accommodate the intuition that we want to know whether people have true beliefs about their sensations.

If the Perceptual Model is formulated using a baseline that tracks our sensations it gives rise to skeptical worries. If ultimate arbiter of whether someone is having a sensation is a private baseline that tracks the actual sensations, how are we to assume anything about people's ability to offer correct introspective reports? With a private baseline we are unable to evaluate the truth-value of any introspective report that we cannot evaluate by the public baseline. If you persist in claiming that you feel a certain way despite whatever evidence I provide, the only evidence I can have in determining that you are cold is hidden from me. I lack the evidence I need to determine whether you are having a sensation because mental states are private and not luminous.

The person introspecting is in no better position to evaluate whether she is a reliable introspector. If the private baseline is something available upon introspection it will inherit the same problems as anything else available to introspection. If a person claims to know through introspection that this is the way she feels when she feels cold because she always feels this way when she feels cold, she will have to give us a reason to think she have applied the baseline correctly... which will in turn require she provide something known through introspection and subject to the same doubts. If the baseline is not something known through introspection, the subject still must use introspection in

applying it. We should then question whether the baseline has been correctly applied. Insofar as our mental states are not luminous, there is nothing to anchor our evaluation of them. We are left with skepticism about mental life.

The Perceptual Model assumes a gap between having a certain sensation and knowing that one is having a sensation. This gap creates a dilemma for the Perceptual Model, either it must use a public baseline to judge all introspective reports, or it must posit a private baseline. If only the public baseline is used, the Perceptual Model will be unable to evaluate cases where there is a disconnect between the sensation and the public expression of that sensation. If the Perceptual Model uses a private baseline, we are forced into skepticism about all mental life. Neither horn of the dilemma seems appealing. Neither Constitutivism nor Expressivism are faced with the dilemma because for them everything in our mental lives is public and evaluable by a public baseline.

6 Conclusion

Williamson's argument initially seems plausible; however, it relies on an inadequately supported premise. Williamson needs to assume that the process by which we come to know our mental states is fallible, but with little reason to do so when there are several other plausible positions. Without assuming a fallible inner sense, the anti-luminosity argument looks like just another sorites argument. Constitutivism and Expressivism's *prima facie* plausibility is bolstered by the Perceptual Model's inability to show that we are reliable introspectors.³⁵

³⁵ I would like to thank Berit Brogaard, who suffered through many drafts of this paper in an attempt to make it better. I would also like to thank John Hawthorne, Richard Lauer, Gualtiero Piccinini, Eric Wiland, and Michael Vossen for reading and commented on previous drafts of this paper.

Bibliography

- Armstrong, D.M. (1968) *A Materialist Theory of Mind*. Routledge: New York, NY.
- Bar-On, Dorit. (2009) "First-Person Authority: Dualism, Constitutivism, and Neo-Expressivism." *Erkenntnis*. 71:1, 53-71.
- (2004) *Speaking My Mind*. Oxford University Press: Oxford, UK.
- Berker, Selim. (2008) "Luminosity Regained." *Philosophers' Imprint*. 8:2, 1-22.
- Byrne, Alex. (2005) "Introspection." *Philosophical Topics*. 33, 79-104.
- Carruthers, Peter. (2009) "How do we know our own minds: The Relationship Between Mindreading and Metacognition." *Behavioral and Brain Sciences*. 2:2, 121-138.
- Dennett, Daniel. (1978) "Why you can't make a computer that feels pain." *Synthese*. 38:3, 415-456.
- Dretske, Fred. (2003) "How Do You Know You Are Not a Zombie?" in Brie Gertler ed. *Privileged Access: Philosophical Accounts of Self-Knowledge*. Ashgate: UK.
- Finkelstein, David H. (2003) *Expression and the Inner*. Harvard University Press: Cambridge, MA.
- Howson, Colin. (1997) "A Logic of Induction." *Philosophy of Science*. 64, 268-90.
- Hurlburt, Russ. T., and Schwitzgebel, Eric. (2007) *Describing Inner Experience? Proponent Meets Skeptic*. Cambridge, MA: MIT Press.
- Hurlburt, Russ. T. (1993) *Sampling inner experience in disturbed affect*. New York: Plenum Press.
- Moran, Richard. (2001) *Authority and Estrangement*. Princeton University Press: Princeton, NJ.
- Neta, Ram and Guy Rohrbaugh. (2004) "Luminosity and the Safety of Knowledge." *Pacific Philosophical Quarterly*. 85, 396-406.
- Nisbett, Richard E., and Timothy Decamp Wilson. (1977) "Telling More Than We Can Know: Verbal Reports on Mental Processes." *Psychological Review*. 84:3, 231-259.
- Piccinini, Gualtiero. (2003) "Data from Introspective Reports: Upgrading from Commonsense to Science." *Journal of Consciousness Studies*. 10.9:10, pp. 141-156.
- (2009) "First-Person Data, Publicity & Self-Measure." *Philosophers' Imprint*. 9:9, 1-6.
- Prinz, Jesse. (2004) "Emotions Embodied." in Robert Solomon ed. *Thinking About Feeling: Contemporary Philosophers on Emotions*. Oxford University Press, Oxford UK.
- Ramachandran, Murali. (2009) "Anti-luminosity: Four Unsuccessful Strategies." *Australasian Journal of Philosophy*. 87:4, 1-14.
- Schacter, Stanley, and Jerome Singer. (1962) "Cognitive, Social and Physiological Determinants of Emotional States." *Psychological Review*. 69, 379-99.
- Shoemaker, Sydney. (1994) "Self-Knowledge and 'Inner Sense.'" *Philosophy and Phenomenological Research*. 54:2, 249-314.
- Tye, Michael. (2002) "Representationalism and the Transparency of Experience." *Nous*. 36:1, 137-151.
- (2009) *Consciousness Revisited: Materialism without Phenomenal Concepts*. MIT Press: New York, NY.
- Williamson, Timothy. (1994) *Vagueness*. Routledge: New York, New York.
- (2000) *Knowledge and its Limits*. Oxford University Press: Oxford, UK.
- Wittgenstein, Ludwig. (1958) *Philosophical Investigations*. G.E.M. Anscombe tr. Prentice Hall: Upper Saddle River, NJ.
- Wright, Crispin. (1998) "Self Knowledge: The Wittgensteinian Legacy." in Crispin Wright, Barry C. Smith, and Cynthia Macdonald eds. *Knowing our Minds*. Oxford University Press: Oxford, UK.