

# OUR COGNITIVE HOMES

## ABSTRACT

Do we bear a special kind of epistemic or cognitive relationship to our core mental states? On a certain Cartesian Conception, such states are *luminous*: if one is in a core mental state then one is in a position to know that one is. Williamson (2000) argues that there are no (non-trivial) luminous conditions. Contra Williamson, Hawthorne (2005) argues that our core mental states are best seen as ‘cozy’: if one is *determinately* in a core mental state then one is in a position to know that one is. Similarly, Conee (2005) proposes that our core mental states are ‘centrally luminous’: if one is *intensely* in a core mental state then one is in a position to know that one is (cf. Sosa 2009). In this paper I show, via a variety of different arguments, that our core mental states are neither luminous, nor cozy, nor centrally luminous. More generally, I show that while we do not bear any special *epistemic* relationship to our core mental states, we nonetheless bear a strong enough *cognitive* relation to them such that there is indeed a theatre of experience which is sufficiently distinctive to comprise a cognitive home.

### *1. Preamble.*

Do we bear a special kind of epistemic or cognitive relationship to our core mental states? On a certain Cartesian conception of the mental (hereafter the ‘Cartesian Conception’), our core mental states (such as *feeling pain*, *feeling cold*, *feeling hot*, and so on) are *luminous*, where a condition C is luminous for subject *s* just in case if C obtains then *s* is in a position to know that C obtains.<sup>1</sup> On the Cartesian Conception, we thus bear a particularly strong epistemic relationship to our core mental states. Is such a view defensible? Williamson (1996, 2000) has argued that there are no (non-trivial) luminous conditions. For that reason, Williamson alleges, we are ‘cognitively homeless’—there is no theatre of thought and experience within which our core mental states obtain and to which we have some special kind of epistemic access. In a recent response, Hawthorne (2005) has granted the soundness of Williamson’s anti-luminosity argument but maintains that core

---

<sup>1</sup> Further candidate luminous conditions include phenomenal conditions of the form *it appears to one that p* (see Williamson 2000, p. 96).

mental states are best seen as ‘cozy’ rather than luminous, where a condition *C* is cozy for a subject *s* just in case if *C* *determinately* obtains then *s* is in a position to know that *C* obtains. If Hawthorne is right then Williamson’s argument misses its proper target and the Cartesian Conception remains in the running. In a similar vein, Conee (2005) has proposed that core mental states are best seen as exhibiting *central luminosity*, rather than luminosity, where a condition *C* is centrally luminous for a subject *s* just in case if *C* centrally (roughly: *intensely*) obtains then *s* is in a position to know that *C* obtains. Moreover, Conee takes central luminosity to be ‘virtually the same idea’ as coziness. Again, in a similar vein, Sosa (2009) has argued that our core mental states are best seen as exhibiting *quasi-luminosity*, where a condition *C* is quasi-luminous for a subject *s* just in case there is a degree to which *s* can be in condition *C*, such that if *s* is in condition *C* to at least that degree, then *s* is in a position to know that they are in condition *C*.<sup>2</sup> Weatherson (2004), meanwhile, has offered a rather different kind route to secure that we are cognitively homed via the idea that the (phenomenal) belief that one is in a core mental state simply constitutes being in that core mental state—hence such a belief constitutes knowledge and so the state is luminous.

The goal of this paper is to show that while we do not bear any special *epistemic* relationship to our core mental states, we do bear a strong enough *cognitive* relation to such states such that there is indeed a theatre of experience and thought which is sufficiently distinctive to comprise a cognitive home. The main theses to be established are: (1) There is a plausible alternative anti-luminosity argument in the offing which does not rely on Williamson’s margin for error principles but on the possibility of ‘first-person agnosticism’. (2) Williamson’s anti-luminosity argument can be straightforwardly strengthened to show that there are no (non-trivial) cozy conditions.<sup>3</sup> (3) A further argument can show that there are no (non-trivial) *warrant-luminous* conditions, where a condition *C* is warrant-luminous just in case in every case in which *C* obtains, one is in a position to form a warranted belief that *C* obtains. (4) We cannot even claim that our core mental states are *warrant-cozy*, where a condition *C* is warrant-cozy just in case in every case in which, determinately, *C* obtains, one is in a position to form a warranted belief that *C* obtains. (5) None of the arguments considered so far can show that there are no (non-trivial) centrally luminous conditions. Hence, centrally luminous conditions are not cozy conditions, contra Conee (2005). (6) Nonetheless, there is an important additional argument, due to Conee, which shows that there are no (non-trivial) centrally luminous conditions. (7) Weatherson’s response to Williamson’s original argument cannot secure the luminosity of our core mental states. (8) The state of feeling cold is such that, very

---

<sup>2</sup> Conee (2005, fn. 18) notes that he first heard of such central/cozy/quasi-luminous conditions from a conversation between Fred Feldman and Jim Pryor. Bryan Frances seems to have been the first to propose the condition of coziness (which he calls ‘definite luminosity’) in unpublished work from late 2002. Greenough (2005, p. 176) proposes that a response-dependent theory of vagueness can be patched up by positing a principle analogous to coziness.

<sup>3</sup> Williamson’s reply to Hawthorne is independent from the arguments presented here (see Williamson 2005, pp. 476-478).

roughly, if one *centrally* (i.e. intensely) feels cold, and one has properly considered the matter, then one believes that one feels cold. (9) This state is also such that, very roughly, if one believes that one feels cold then one is not centrally in the state of not feeling cold. (10) Having a cognitive home in the sense argued for is perfectly compatible with a rejection of the phenomenal conception of evidence.

## 2. Williamson's anti-luminosity argument.

The thumbnail version of Williamson's anti-luminosity argument is as follows:<sup>4</sup>

The main idea behind the argument against luminosity is that our powers of discrimination are limited. If we are in a case  $\alpha$ , and a case  $\alpha'$  is close enough to  $\alpha$ , then for all we know we are in  $\alpha'$ . Thus what we are in a position to know in  $\alpha$  is still true in  $\alpha'$ . Consequently a luminous condition obtains in  $\alpha$  only if it also obtains in  $\alpha'$ , for it obtains in  $\alpha$  only if we are in a position to know that it obtains in  $\alpha$ . In other words, a luminous condition obtains in any case close enough to cases in which it obtains (2000, p. 13).

We can unpack this argument as follows: Firstly, consider a gradual phenomenal transition whereby a subject  $s$  knows that they feel hot at time  $t$ , but does not feel hot at time  $t$  plus one hour, where at every state of this transition the subject is in a position to wonder whether or not they feel hot. At  $t$ ,  $s$  forms the belief that they feel hot; at the end of the process,  $s$  forms the belief that they do not feel hot. Secondly, given our limited powers of discrimination, then knowledge of the obtaining of a condition  $C$  (e.g. feeling hot) requires what Williamson calls 'a margin for error' for knowledge, as follows:

(KM) For all cases  $\alpha$ ,  $\beta$ , if in  $\alpha$   $s$  is in a position to know that  $C$  obtains then  $C$  obtains in  $\beta$ , (where  $\beta$  is close to  $\alpha$ ).<sup>5</sup>

Thirdly, suppose that  $C$  is 'luminous':

(KL) For all cases  $\alpha$ , if  $C$  obtains in  $\alpha$  then in  $\alpha$   $s$  is in a position to know that  $C$  obtains.

(see Williamson 2000 p. 95). Fourthly, KM plus KL entail the following soritical principle:

---

<sup>4</sup> The more subtle version of Williamson's argument (2000 pp. 96-98) applies reliability considerations to degrees of confidence. This argument can also be strengthened to show that there are no (non-trivial) cozy conditions. For simplicity, I focus on the less subtle version.

<sup>5</sup> Cases can be thought of as comprising a subject,  $s$ , a time, and a possible world. The quantifiers should be taken to range over physically and psychologically feasible cases for normally functioning human subjects.

(SR) For all cases  $\alpha$ ,  $\beta$ , if in  $\alpha$   $s$  is in a position to know that  $C$  obtains then in  $\beta$   $s$  is in a position to know that  $C$  obtains, (where  $\beta$  is close to  $\alpha$ ).

Given SR, and given that  $s$  is in a position to know that they feel hot at time  $t$  then  $s$  is in a position to know that they feel hot in all cases—in particular at time  $t$  plus one hour. But it is given that  $s$  does not feel hot at time  $t$  plus one hour and so  $s$  is not in a position to know that they feel hot at time  $t$  plus one hour. Contradiction. Upshot: the condition of *feeling hot* is not luminous. Likewise for all other core mental states. Thus, the Cartesian Conception, under which these core mental states are taken to be luminous, must be abandoned.<sup>6</sup> Indeed, the only luminous conditions are trivially luminous conditions.<sup>7</sup>

There are four points of note. Firstly, the argument does not essentially depend on taking ‘feels cold’, ‘feels hot’, etc., or even ‘knows’, to be vague predicates (see Williamson 2000 pp. 102-106, and below). Secondly, the argument is designed to show that luminosity fails even when conditions of judgment are, in the relevant sense, normal (see §11). Thirdly, Williamson takes his margin for error principle KM to be derived from the following safety principle on knowledge:

(KS) For all cases  $\alpha$ ,  $\beta$ , if in  $\alpha$   $s$  is in a position to know that  $C$  obtains, then in  $\beta$  it is not the case that:  $s$  believes that  $C$  obtains and it is not the case that  $C$  obtains, (where  $\beta$  is close to  $\alpha$ ),

plus the doxastic principle:

(B) If  $s$  has a belief in a case  $\alpha$  then  $s$  could easily still have had this belief in a close (but distinct) case  $\beta$  (see Williamson 2000, pp. 126-129).<sup>8</sup>

In what follows, I will not question either KS or B, nor the derivation of KM from KS and B.<sup>9, 10</sup>

Fourthly, the anti-luminosity argument can also be employed to show that there are no (non-trivial) ‘negative’ core mental states. This is because being in a position to know that these states obtain also requires a margin for error. So, the conditions of *not feeling pain*, *not feeling cold*, and so on, are also non-luminous.

---

<sup>6</sup> Shoemaker (1996, ch. 3) defends a form of the Cartesian Conception. I shall not pursue whether Descartes ever advanced such a conception himself.

<sup>7</sup> A trivially luminous condition is a condition which obtains in no case (e.g. impossible conditions) or every case (e.g. necessary conditions).

<sup>8</sup> More refined formulations of KS and B index ‘know’ and ‘believe’ to a method or basis. The core method we are interested in with respect to a belief that a core mental state obtains is introspection (see below).

<sup>9</sup> KS is questioned by Brueckner and Fiocco (2002) and Neta and Rohrbaugh (2004). Cohen, ms, effectively doubts B. Steup (forthcoming) doubts a principle (very loosely) related to B (see fn. 22 below).

<sup>10</sup> KM straightforwardly entails KS (without using B).

### 3. *The quick anti-luminosity argument.*

Suppose that the Cartesian Conception *should* enjoin the thesis that negative core mental states are luminous. It is also possible to show that this augmented conception is flawed by a rather different style of anti-luminosity argument which makes no (overt) appeal to reliability considerations or to a margin for error for knowledge. Firstly, KL entails the following principle of *omniscience*:

(KO) For all cases  $\alpha$ , if C obtains in  $\alpha$  (and  $s$  has actively wondered whether or not C obtains) then in  $\alpha$   $s$  knows that C obtains.

Secondly, given that knowledge entails belief, KO entails the following principle of *credulity*:

(CR) For all cases  $\alpha$ , if C obtains in  $\alpha$  (and  $s$  has actively wondered whether or not C obtains) then in  $\alpha$   $s$  believes that C obtains.

Thirdly, since under the augmented Cartesian Conception in hand ‘C’ ranges over both positive and negative core mental states, then we have the following instances of CR:

(CR1) For all cases  $\alpha$ , if  $s$  feels hot in  $\alpha$  (and  $s$  has actively wondered whether or not they feel hot) then in  $\alpha$   $s$  believes that they feel hot.

(CR2) For all cases  $\alpha$ , if  $s$  does not feel hot in  $\alpha$  (and  $s$  has actively wondered whether or not they feel hot) then in  $\alpha$   $s$  believes that they do not feel hot.<sup>11</sup>

However these principles are far from obviously true. Consider again the gradual phenomenal transition whereby  $s$  feels hot at time  $t$  and does not feel hot at time  $t$  plus one hour, such that at every stage of this transition  $s$  has wondered whether or not they feel hot. There are stages midway through the transition whereby no belief either way need be triggered— $s$  neither feels hot nor feels not hot despite having attentively and fully considered the matter. It’s seems perfectly acceptable for an alert, normally functioning, and competent subject to fail to reach a verdict and so withhold belief as to whether or not they feel hot. Given CR1 and CR2, a contradiction follows. Equally, consider a gradual and slow onset of toothache. At each stage in the process one considers whether or not one feels pain. It seems perfectly acceptable for an alert, normally functioning, and competent subject, at certain stages midway in the transition, to refrain from forming a belief as to

---

<sup>11</sup> In Shoemaker’s terminology (1996, p. 51), the state of feeling hot is strongly self-intimating if and only if it satisfies both CR1 and CR2, and weakly self-intimating if and only if it satisfies CR1. Shoemaker takes core mental states to be strongly self-intimating.

whether or not they are in pain—one might feel a faint throbbing in one’s tooth and yet feel unsure that this feeling amounts to a feeling of being in pain and unsure that this feeling does not amount to feeling in pain.

Doubtless, many competent subjects, indeed perhaps most subjects, may not express such doxastic hesitancy. However, all that is required to get the anti-luminosity argument in hand to work is for there to be one subject, who is fully alert, fully competent with the concepts in question, who is functioning normally, who is given sufficient time to reflect as to whether or not they are in the core mental state, and yet refrains from forming a belief as to whether or not the core mental state obtains when that state does/does not obtain. This style of argument involves no overt appeal to reliability considerations or to a margin for error. Rather, the quick argument utilises a form of what Shoemaker (1996, p. 51) calls ‘first-person agnosticism’.<sup>12</sup> Should there be doubts about the soundness of Williamson’s own anti-luminosity argument, perhaps via doubts about KS or B, or doubts about the derivation of KM from KS and B, then this quick argument is arguably much harder to undermine.<sup>13</sup>

#### 4. *Cozy but non-luminous cognitive homes.*

Hawthorne (2005) accepts the soundness of Williamson’s anti-luminosity argument. However, he maintains that this argument misses its proper target because our core mental states are best seen as ‘cozy’ rather than luminous, where a condition C is cozy if and only if the following ‘natural fall-back’ principle is valid:

(KDL) For all cases  $\alpha$ , if it is *determinately* the case that C obtains in  $\alpha$  then in  $\alpha$  *s* is in a position to know that C obtains.<sup>14</sup>

Crucially, KM plus KDL does not entail SR. Furthermore, all that an analogue of the quick argument can show is that when a competent, normally functioning subject, after proper

---

<sup>12</sup> Reliability considerations may play a covert role in that epistemically successful subjects may have evolved to fail to form beliefs where there is a danger that their beliefs may be false.

<sup>13</sup> Shoemaker finds such agnosticism plausible for certain mental states (such as whether the subject has an Oedipus Complex) but thinks that such agnosticism is totally implausible when it comes to ‘sensations and ordinary beliefs and desires’. It is clear, however, that Shoemaker is not considering cases where a sensation or core mental state obtains, but only very faintly. In such cases, first-person agnosticism is surely very plausible, hence the quick argument. Cf. Greenough (2005, p. 175) on ‘the problem of silence’.

<sup>14</sup> See Hawthorne (2005, p. 476). KL entails KDL given that if C determinately obtains then C obtains. Hawthorne also offers a refined analysis of coziness as a property of concepts and not conditions: ‘a concept C, denoting c, is cozy if and only if in any case in which C determinately holds, one is in a position to know that c obtains by an exercise of C’ (2005, pp. 453–454). For simplicity, we can ignore this (important) complication in what follows.

consideration, fails to believe that they feel cold (at time  $t$ ) and fails to believe that they do not feel cold (at  $t$ ) then, at  $t$ , it is simply indeterminate whether  $C$  obtains.

The point of replacing KL with KDL is to avoid any worries about vagueness. Here the thought is that one barrier to being in a position to know that  $C$  obtains might simply be that it is borderline that  $C$  obtains, where, generically, this means that  $C$  neither determinately obtains nor determinately fails to obtain. Since the predicates ‘feels pain’, ‘feels hot’, ‘feels cold’, and so on, are all surely vague, and thus give rise to borderline cases, everybody should accept that KL fails for reasons of vagueness. What the anti-Cartesian needs to establish is that KL fails for reasons other than the presence of vagueness. Such is Hawthorne’s challenge.

### 5. *Determinacy.*

Is ‘determinately’ an epistemic or a non-epistemic operator? In Greenough (2003, p. 252) it is argued that, from a theory-neutral perspective, ‘definitely  $F$ ’, and its cognate ‘determinately  $F$ ’, are best taken to mean ‘known to be  $F$ ’.<sup>15</sup> So, to say that  $C$  neither determinately obtains nor determinately fails to obtain is just to say that it is neither known that  $C$  obtains nor known that  $C$  fails to obtain. On that reading, KDL is trivially true for any condition whatsoever. If KDL is to bear the weight of the Cartesian Conception, ‘determinately’ must receive a non-epistemic reading. (This in any case seems to be Hawthorne’s intended reading.)

On a standard supervenient conception of vagueness, ‘determinately’ is used to express the notion of supertruth in the object language as follows: ‘Determinately  $A$ ’ is true if and only if ‘ $A$ ’ is supertrue, that is, true in all admissible sharpenings. On this conception, truth in natural language just is determinate truth/supertruth (see Fine 1975, Keefe 2000, and, for critical discussion, Williamson 1994, p. 149). On a Łukasiewicz three-valued conception of vagueness, ‘determinately’ can be used to express a strong notion of truth in the object language (whereby if ‘ $A$ ’ is gappy then ‘It is (strongly) true that  $A$ ’ is false) as follows: ‘It is determinately true that  $A$ ’ is true if and only if ‘ $A$ ’ is (strongly) true. On a continuum many-valued conception, ‘determinately’ can be used to express the notion of truth to degree 1 in the object language as follows: ‘It is determinately true that  $A$ ’ is true if and only if ‘ $A$ ’ is true to degree 1.

While these different conceptions of vagueness/determinacy each yield a different reading of KDL, they nonetheless all entail the following principle:

---

<sup>15</sup> Cf. Williamson (1994, pp. 194-195, 1995).

(K) For all cases  $\alpha$ , if in  $\alpha$   $s$  is in a position to know that  $C$  obtains/does not obtain then it is determinately the case that  $C$  obtains/does not obtain in  $\alpha$ .

K is typically invoked to explain why, given that a condition  $C$  neither determinately obtains nor determinately fails to obtain, a subject is in no position to know whether or not  $C$  obtains. If it is not true that  $C$  obtains, or if it is not true to degree 1 that  $C$  obtains, then this precludes knowledge that  $C$  obtains.<sup>16</sup> We can take this fact into account in specifying a stronger safety condition on knowledge.

#### 6. *Determinate Safety, Determinate Margins for Error, and Anti-Coziness.*

Given K, KS should be replaced with the following principle of *Determinate Safety*:

(KDS) For all cases  $\alpha$ ,  $\beta$ , if in  $\alpha$   $s$  is in a position to know that  $C$  obtains, then in  $\beta$  it is not the case that:  $s$  believes that  $C$  obtains and it is not *determinately* the case that  $C$  obtains, (where  $\beta$  is close to  $\alpha$ ).

The standard supervaluationist reads KDS as saying that one is in a position to know that  $C$  obtains only if one's belief that  $C$  obtains couldn't easily have been *untrue*. Likewise for the defender of a Łukasiewicz three-valued conception of vagueness. On a continuum many-valued approach, KDS is read as saying that one is in a position to know that  $C$  obtains only if one's belief that  $C$  obtains couldn't easily have been less than true to degree 1. In general, on all non-epistemic accounts of vagueness under which K is valid, if one's belief could easily have been indeterminate in truth-value (and so could easily have been not determinately true) then that belief cannot constitute knowledge. Given KDS, one can then derive the following margin for error principle:

(KDM) For all cases  $\alpha$ ,  $\beta$ , if in  $\alpha$   $s$  is in a position to know that  $C$  obtains then  $C$  determinately obtains in  $\beta$  (where  $\beta$  is close to  $\alpha$ ).<sup>17</sup>

But now we have the materials for a strengthened anti-coziness argument because KDM plus KDL *does* entail SR. So, insofar as Williamson's original anti-luminosity argument is sound then the anti-coziness argument just given is sound also.<sup>18</sup>

---

<sup>16</sup> On certain non-epistemic conceptions of indeterminacy under which classical logic and classical semantics remain valid, truth is not determinate truth (see e.g. McGee and McLaughlin 1995, Greenough 2008). Even so, on these conceptions principle K remains valid.

<sup>17</sup> Given the doxastic principle B.

<sup>18</sup> KDS entails KS, and KDM entails KM, given that if  $C$  determinately obtains then  $C$  obtains.

Are there any weaker luminosity principles which are immune from a Williamsonian style anti-luminosity argument, immune from the quick argument, and yet sufficiently strong to sustain the something akin to the Cartesian Conception?

### 7. Warrant-luminosity.

One further fallback luminosity principle exploits the notion of *being in a position to have a warranted belief* rather than the notion of being in a position to know. Here the idea is that while our limited powers of discrimination undermine both luminosity and coziness they do not undermine the following weaker principle of *warrant-luminosity*:

(WL) For all cases  $\alpha$ , if C obtains in  $\alpha$  then in  $\alpha$   $s$  is in a position to have a warranted belief that C obtains.

The notion of being in a position to have a warranted belief invoked here is taken to be non-factive.<sup>19</sup> Let this notion also be such that being in a position to warrantably believe that C obtains requires *semi-safety*, where  $s$ 's belief that C obtains is semi-safe if and only if in *nearly all* nearby worlds it is not the case that:  $s$  believes that C obtains and C does not obtain. And so, we have the following semi-safety condition on being in a position to have a warranted belief, call this *warrant-safety*:

(WS) For all cases  $\alpha$ , if in  $\alpha$   $s$  is in a position to have a warranted belief that C obtains, then in nearly all cases  $\beta$  it is not the case that:  $s$  believes that C obtains and it is not the case that C obtains, (where  $\beta$  is close to  $\alpha$ ).<sup>20</sup>

The corresponding margin for error principle is then:

(WM) For all cases  $\alpha$ , if in  $\alpha$   $s$  is in a position to have a warranted belief that C obtains, then in nearly all cases  $\beta$  C obtains in  $\beta$ , (where  $\beta$  is close to  $\alpha$ ).

Crucially, WM plus WL do not entail SR. So, warrant-luminosity looks to be in good standing and a weakened form of the Cartesian Conception remains in the running.

---

<sup>19</sup> KL entails WL given that being in a position to know requires being in a position to have a warranted belief.

<sup>20</sup> This notion of warrant is such that while one cannot have a warranted belief that one is standing before a barn (in the standard Goldman-Ginet barn-façade case where there are many barn-façades in the vicinity), one can nonetheless have a warranted belief that one's ticket is not the winning ticket in a fair lottery (and so warrant does not satisfy so-called multi-premise closure).

## 8. *Against warrant-luminosity.*

The trouble with this gambit is that while being in a position to have a warranted belief is non-factive, this position is such as to nonetheless satisfy the D-axiom of standard modal logics such that:

(D) It is not the case that: in  $\alpha$   $s$  is in a position to have a warranted belief that  $C$  obtains and in  $\alpha$   $s$  is in a position to have a warranted belief that  $C$  does not obtain.

Any non-factive state that satisfies a principle like D (for example, warranted belief, reasonable belief, probabilified belief, and the ‘being in a position to’ variants of such states) will nonetheless satisfy a weaker margin for error principle. In the case of being in a position to warrantably believe, this principle is:

(WWM) For all cases  $\alpha, \beta$ , if in  $\alpha$   $s$  is in a position to have a warranted belief that  $C$  obtains, then in  $\beta$  it is not the case that  $s$  is in a position to have a warranted belief that  $C$  fails to obtain (where  $\beta$  is close to  $\alpha$ ),

which is classically equivalent to:

(WWM) There is no case  $\alpha, \beta$ , such that in  $\alpha$   $s$  is in a position to have a warranted belief that  $C$  obtains and in  $\beta$   $s$  is in a position to have a warranted belief that  $C$  fails to obtain (where  $\beta$  is close to  $\alpha$ ).

WWM is designed to capture the plausible idea that our powers of discrimination are limited in the following sense: if one is in a position to warrantably believe that one is in a case  $\alpha$ , and a case  $\alpha'$  is close enough to  $\alpha$ , then for all one is in a position to warrantably believe, one is in  $\alpha'$ .<sup>21</sup>

Now we have all the materials for an argument which shows that WWM plus WL entails a contradiction (at least if WL is taken to hold for both positive negative core mental states). Let the relevant problematic instances of WL be:

(WL1) For all cases  $\alpha$ , if  $s$  feels hot in  $\alpha$  then in  $\alpha$   $s$  is in a position to have a warranted belief that they feel hot.

(WL2) For all cases  $\alpha$ , if  $s$  does not feel hot in  $\alpha$  then in  $\alpha$   $s$  is in a position to have a warranted belief that they do not feel hot.

WWM, as applied to *feeling hot*, yields:

---

<sup>21</sup> WM entails WWM though not vice versa.

(WWM1) There is no case  $\alpha, \beta$ , such that in  $\alpha$   $s$  is in a position to have a warranted belief that they feel hot and in  $\beta$   $s$  is in a position to have a warranted belief that they do not feel hot (where  $\beta$  is close to  $\alpha$ ).

Given WL1 and WL2, WWM1 entails:

(SR1) There is no case  $\alpha, \beta$ , such that in  $\alpha$   $s$  feels hot and in  $\beta$   $s$  does not feel hot (where  $\beta$  is close to  $\alpha$ ).

But SR1 entails that  $s$  feels hot in all cases or in no case. Given that  $s$  feels hot in some but not all cases then a contradiction results. So, the conjunction of WL1 and WL2, and thus WL, must be abandoned. Upshot: there are no (non-trivial) core mental states which satisfy warrant-luminosity.

There are six points of note: Firstly, a general moral emerges. Any notion of warrant or justification that is sufficiently interesting to sustain the broad motivation to accept the Cartesian Conception will always be strong enough to satisfy some weak margin for error principle along the lines of WWM. So, any luminosity principle along the lines of WL is destined to fail.

Secondly, an analogue of the quick argument runs as follows: WL entails a principle of *warrant-omniscience* to the effect that:

(WO) For all cases  $\alpha$ , if  $C$  obtains in  $\alpha$  (and  $s$  has actively wondered whether or not  $C$  obtains) then in  $\alpha$   $s$  has a warranted belief that  $C$  obtains.

WO, like KO, entails the principle of credulity CR. Given that a competent, fully alert, normally functioning subject, after proper consideration, can both fail to believe that they feel cold and fail to believe that they do not feel cold (at a single time) then a contradiction follows.

Thirdly, Williamson's margin for error principle KM is a much stronger principle of limited discrimination than WWM. Suppose one doubts that knowledge requires safety. Given that KM entails KS, one must therefore doubt KM too. But such doubts need not trouble WWM.

Fourthly, given that being in a position to know entails being in a position to warrantably believe, WWM entails the following weak principle of limited discrimination:

(KWM) There is no case  $\alpha, \beta$ , such that in  $\alpha$   $s$  is in a position to know that  $C$  obtains and in  $\beta$   $s$  is in a position to know that  $C$  fails to obtain (where  $\beta$  is close to  $\alpha$ ).

KWM is designed to capture the plausible idea that our powers of discrimination are limited in the following sense: if one is in a position to know that one is in a case  $\alpha$ , and a case  $\alpha'$  is close enough to  $\alpha$ , then for all one is in a position to know, one is in  $\alpha'$ . Given the factivity of being in a position

to know, KM entails KWM but not vice versa. Indeed, given Williamson's (2000, ch.11) knowledge account of assertion, and the attendant knowledge account of belief whereby one is in a position to warrantably believe that C obtains if and only if one is in a position to know that C obtains, then WWM and KWM are effectively equivalent. Under that account, KM entails WWM, but not vice versa. One could, for example, accept Williamson's knowledge account of assertion and belief, under which KWM and WWM are equivalent, and yet hold that knowledge does not require safety but merely semi-safety. On such a position, KM is invalid but KWM remains valid.

Fifthly, KWM can be used to generate a further anti-luminosity argument analogous to the argument given above: just as WWM plus WL entails a contradiction, KWM plus KL entails a contradiction (at least if KL holds good for both positive and negative conditions).<sup>22</sup>

Sixthly, it may be felt that using WWM to show that WL fails is just question-begging. (Likewise, for using KWM to show that KL fails.) The thought here is that to simply assume that our powers of discrimination are limited, in the sense given by WWM, is just to assume that WL is false. (Likewise for using KWM to show that KL fails). Hence, the arguments using WWM and KWM are unpersuasive. Let us assume that there is at least some force to this worry. The result is then stalemate. The Cartesian assumes that WL/KL are valid, the anti-Cartesian assumes that WWM/KWM are valid. Can the deadlock be broken?

There are three ways in which this might be done. Firstly, the Cartesian must also grant that 'feels hot', 'feels cold', and so on, are vague terms (or that the properties expressed by these predicates are vague properties). All partisans to the vagueness debate can agree that it is a minimal symptom of the presence of vagueness that vague terms draw no known boundary across their associated dimension of comparison (see Greenough 2003). But that effectively sanctions KWM. And since WWM is more or less equi-plausible with KWM, we get WWM back too. The Cartesian must then find some way of arguing that it is not a minimal symptom of vagueness that vague terms draw no known boundary across their associated dimension of comparison. The prospects seem dim.<sup>23</sup> Secondly, the Cartesian position should not be the default view as principles like KL and WL are *exceedingly* strong claims, these principles do, after all, entail a kind of local omniscience

---

<sup>22</sup> Steup (2009) effectively reconstructs Williamson's derivation of KL from KS as proceeding via the following principle (rather than via B): for all cases  $\alpha, \beta$ , if in  $\alpha$   $s$  knows that C obtains then in  $\beta$   $s$  believes that C obtains (if  $s$  has wondered whether C obtains, and where  $\beta$  is close to  $\alpha$ ). Steup calls this principle 'PR2'. But if PR2 has any plausibility then so does the following (weaker) principle: for all cases  $\alpha, \beta$ , if in  $\alpha$   $s$  knows that C obtains then in  $\beta$   $s$  does not believe that C does not obtain (if  $s$  has wondered whether C obtains, and where  $\beta$  is close to  $\alpha$ ). But given that failing to believe that C does not obtain entails failing to know that C does not obtain, then this weaker principle entails the KWM-like principle: there is no case  $\alpha, \beta$ , such that in  $\alpha$   $s$  knows that C obtains and in  $\beta$   $s$  knows that C fails to obtain (where  $\beta$  is close to  $\alpha$ ). But this KWM-like principle, *on its own*, undermines principle KO, and so principle KL (at least if 'C' is taken to range over both positive and negative conditions). So, Williamson should resist Steup's reconstruction of his derivation of KL from KS as proceeding via PR2.

<sup>23</sup> Moreover, even if some theory of vagueness does entail that this feature is not a symptom of vagueness, the Cartesian Conception (based on WL and WDL) is then only defensible from within the perspective of some idiosyncratic theory of vagueness. As such, it will have lost its general appeal.

(principles KO and WO respectively). Indeed, the Cartesian must grant that introspection is less than perfectly discriminatory for non-core mental states, so we then need an *argument* to show why introspection suddenly (and magically) becomes perfectly discriminatory with respect to our core mental states. In the absence of such an argument we are entitled to assume (from the outset at least) that principles WWM and KWM hold good for introspection across the board. The third point is that even if the stalemate cannot be broken by the considerations just given the quick argument against WL and KL remains intact. Thus the Cartesian must yield in any case.

### 9. *Warrant-coziness.*

A further Cartesian gambit might be to replace WL with its ‘cozy’ variant, yielding a principle of *warrant-coziness*:

(WDL) For all cases  $\alpha$ , if C determinately obtains in  $\alpha$  then in  $\alpha$   $s$  is in a position to have a warranted belief that C obtains.

As it turns out, WDL is compatible with WWM.<sup>24</sup> Indeed, just as with KDL, all that an analogue of the quick argument can show is that when a fully alert, normally functioning, competent subject, after proper consideration, fails to form a belief that they feel cold and fails to form a belief that they do not feel cold (at a single time) then it is indeterminate whether C obtains (at that time).

However, this gambit is dialectically suspect. The reason is that from the perspective of a theory which takes ‘feels hot’, ‘feels cold’, and so on, to draw sharp boundaries, (such as a standard epistemicist theory of vagueness) there is no difference between WL and WDL. Given such a theory, if condition C obtains then condition C determinately obtains, and vice versa, (on each of the non-epistemic reading of ‘determinately’ mooted above). Given this conditional, WDL entails WL, and since WL fails then so must WDL. While it may not in the end turn out that ‘feels hot’ draws a sharp boundary, it would be very odd if the viability of the Cartesian Conception turned on whether or not this was so. Since the correct dialectical stance within the current setting ought to be neutrality as to whether or not ‘feels hot’, ‘feels cold’, and so on, draw sharp boundaries then we cannot claim that our core mental states are semi-cozy. Some alternative conception must be sought which is free from such dialectical difficulties.<sup>25</sup>

---

<sup>24</sup> Likewise, KDL is compatible with KWM.

<sup>25</sup> Such a dialectical point could also have been made earlier against Hawthorne’s suggestion of replacing KL with KDL since, for the epistemicist, such principles are equivalent. However, given KDM, a much stronger objection was available against KDL.

## 10. Central luminosity and quasi-luminosity.

In a somewhat similar vein to Hawthorne, Conee (2005) has proposed that core mental states are best seen as exhibiting what may be termed *central luminosity*, rather than luminosity. And so KL is to be replaced with KCL:

(KCL) For all cases  $\alpha$ , if C centrally (i.e. *intensely*) obtains in  $\alpha$  then in  $\alpha$   $s$  is in a position to know that C obtains.

With respect to feeling hot, the basic idea is that if a subject *intensely* feels hot then they are at least in a position to know that they feel hot. Conee (2005, fn. 18) takes coziness and centrality to be ‘virtually the same idea’. If that is so then the arguments from §6 can simply be reproduced to show that there are no (non-trivial) centrally luminous conditions. However, central luminosity, though superficially similar to coziness, is not (virtually) the same idea at all. From the perspective of an epistemicist conception of vagueness, if C obtains then C determinately obtains. However, from that perspective, it is not the case that if C obtains then C centrally obtains. Indeed, with respect to all plausible theories of vagueness, if one is in a position to know that C obtains then C determinately obtains (i.e. principle K is valid). However, it is not the case that if one is in a position to know that C obtains then C centrally obtains. This explains why there are no centrally luminous versions of KS and KM, and hence no strengthened anti-luminosity argument against KCL.<sup>26</sup>

Equally, the following principle of central luminosity with respect to warrant is compatible with WWM:

(WCL) For all cases  $\alpha$ , if C centrally (i.e. *intensely*) obtains in  $\alpha$  then in  $\alpha$   $s$  is in a position to have a warranted belief that C obtains.

Furthermore, all that an analogue of the quick argument against KCL and WCL can show is that when a competent subject, after proper consideration, fails to believe that they feel cold and fails to believe that they do not feel cold (at a single time) then C neither centrally obtains nor centrally fails to obtain (at that time).

In a similar vein to Conee, Sosa (2009) has argued that our core mental states are best seen as exhibiting what he calls ‘quasi-luminosity’, where a condition is quasi-luminous just in case ‘if

---

<sup>26</sup> KDL entails KCL (but not vice versa) given that if C centrally obtains then C determinately obtains. Given that KL entails KDL (but not vice versa), then KL entails KCL (but not vice versa). KCL entails WCL given that being in a position to know requires being in a position to have a warranted belief. And so, KL entails WCL.

there is a degree to which one can be in it, such that if one is in it at least to that degree, then one is in a position to know that one is in it', which we can restate as follows:

(KQL) For some degree  $d$  and for all cases  $\alpha$ , if in  $\alpha$   $C$  obtains to at least degree  $d$  then in  $\alpha$   $s$  is in a position to know that  $C$  obtains.<sup>27</sup>

Suppose condition  $C$  intensely obtains just in case  $C$  obtains to at least degree  $d_1$ . Principle KCL then becomes: for all cases  $\alpha$ , if in  $\alpha$   $C$  obtains to at least degree  $d_1$  then in  $\alpha$   $s$  is in a position to know that  $C$  obtains. Given existential generalisation, KQL follows. Likewise, if condition  $C$  determinately obtains just in case  $C$  obtains to degree  $d_2$ , then KDL entails KQL. Sosa does not say whether KQL is motivated by the idea that our core mental states are cozy or by the idea that they are centrally luminous. If the former, then, given KDM and the attendant anti-coziness argument, KQL is unmotivated. If the latter, then KQL is well motivated but for all intents and purposes we can replace KQL with KCL (as I shall do so in what follows).

It thus appears that we have hit upon one or more principles which are immune from any of the anti-luminosity style arguments we have considered so far. Furthermore, while there is a dialectical objection against replacing WL with WDL (and a corresponding dialectical objection against replacing KL with KDL), there is no dialectical objection against replacing KL with KCL (and replacing WL with WCL). This is because, *properly construed*, the rationale behind positing that core mental states are centrally luminous is not to avoid any worries concerning vagueness, but to offer a set of principles which can both accommodate the fact that introspection is less than perfectly discriminatory and yet be sufficiently strong to sustain a broadly Cartesian conception of the mental.

However, there is one further important species of anti-luminosity argument yet to be considered.

### *11. Conee's anti-luminosity argument.*

Williamson's anti-luminosity argument proceeds under the generous (implicit) assumption that conditions are normal/good both internally and externally for the subject. In the external case, this assumption includes the claim that there are no defeaters of the subject's grounds for their belief as to whether or not they are in some core mental state. Conee (2005, pp. 48-9) drops this external

---

<sup>27</sup> KQL only applies to conditions which come in degrees. I set aside the issue as to whether all candidate (core) mental states come in degrees. Arguably, *feeling hot* does come in degrees.

element of Williamson's assumption and runs the following *defeasibility* anti-luminosity argument (where I extrapolate somewhat from Conee's own presentation):

Suppose we accept something like the following defeasibility condition on being in a position to have a warranted belief/know that C obtains:

(1) If in  $\alpha$   $s$  has a strong unoverridden reason to doubt that C obtains then in  $\alpha$   $s$  is not in a position to have a warranted belief that C obtains (and so is not in a position to know that C obtains given that knowing entails being warranted).

Suppose that:

(2) In  $\alpha$ ,  $s$  is in the state of centrally (i.e. *intensely*) feeling cold and so in the state of feeling cold.

Suppose further that:

(3) In  $\alpha$ , the following (undefeated) defeater obtains: a person who  $s$  fully trusts and takes to be a world expert on the nature and causes of phenomenal states, is in a position to tell  $s$  that they are not functioning normally due to some rare brain damage, acquired at birth, such that, under the present conditions, they have a *delusional* feeling that they feel cold when in fact they do not really feel cold at all.

Given that the obtaining of this defeater constitutes a strong unoverridden reason to doubt that  $s$  feels cold, then:

(4) Despite the fact that in  $\alpha$ ,  $s$  is in the phenomenal state of centrally feeling cold,  $s$  is not in a position to have a warranted belief that they feel cold and so  $s$  is not in a position to know that they feel cold.

There are three points of note. Firstly, the version of Conee's argument just given is thoroughly externalist because the mere fact that the trusted, 'expert' is *in a position* to offer their testimony is enough to defeat the subject's epistemic position. The subject need not hear this testimony in order for the evidential defeat to occur. Here a defeater is an *undercover fact*—a fact of which  $s$  is presently unaware, but is such that  $s$  could easily become aware of it, such that, were  $s$  to become aware of it, it would undermine  $s$ 's present evidence for believing that they feel cold. On this externalist reading, two subjects can be in identical phenomenal states and yet one subject is warranted in believing that they feel cold, yet the other is not so warranted. Hence, the argument rules out what Williamson (2000, p. 173) calls 'the phenomenal conception of evidence'.

Secondly, the argument undermines all of the Cartesian principles we have considered so far (i.e. KL, KDL, KO, WL, WDL, WO, KCL and WCL). However, given the externalist nature of the

argument, it cannot undermine certain thoroughly internalist versions of luminosity, coziness, and central luminosity. Given the possibility of the quick argument, and given the dialectical illegitimacy of denying that if C obtains then C determinately obtains, then the most promising of these internalist principles is as follows:

(JCL) For all cases  $\alpha$ , if C centrally (i.e. *intensely*) obtains in  $\alpha$  then in  $\alpha$   $s$  is in a position to justifiably believe that C obtains.

This principle is untouched by Conee's argument (as presented above) because a defeater (standardly conceived) does not undermine one's justification for belief, even though it can undermine being in a position to know/have a warranted belief. Let the notion of being in a position to justifiably believe invoked here be such that a subject who is a bodiless but sentient brain-in-a-vat lacks knowledge that they have two hands, and indeed lacks warrant to believe that they have two hands, but is nonetheless justified in believing that they have two hands on the basis that (a) it appears that they have two hands and (b) that they are justified in believing that appearances are veridical).<sup>28</sup>

It looks like JCL is the last chance saloon for a broadly Cartesian conception of the mental under which we bear a special epistemic relationship to our core mental states. As it turns out, a more internalist version of Conee's argument is equipped to undermine JCL. Suppose that the subject in fact hears the defeating testimony. In such a case, there remains *some* justification to believe that they feel cold because the original phenomenal justification remains in place: it still appears that they feel cold. However, due to the given testimony, they now lack justification to believe that such appearances are veridical and so they are not in a position to justifiably believe that they feel cold. So, it looks like the following principle is all that remains of a broadly Cartesian view:

(JCL)\* For all cases  $\alpha$ , if C centrally obtains in  $\alpha$  then in  $\alpha$   $s$  is in a position to have some justification to believe that C obtains.

So, if C centrally obtains and  $s$  fully considers whether C obtains then  $s$  will have *some* justification to believe that C obtains—but, as we have seen, that is compatible with having some justification for believing that C does not obtain. Clearly JCL\* is far too weak to sustain anything like the original motivation behind the Cartesian Conception.

Thirdly, it might be felt that the mere fact that a subject feels pain, for example, is always such as to defeat any potential defeater of the subject's warrant/justification for believing that they feel

---

<sup>28</sup> Here I am assuming the notion of warrant set out in §7.

pain. As a result, principles KL *et al.* are untouched by Conee's argument.<sup>29</sup> In the first place, this doesn't seem at all plausible when a subject feels in pain, *but only very faintly*, and forms the justified true belief that they feel pain on the basis of their awareness that they do (plus the background belief that the deliverances of their awareness are veridical). Here let the countervailing trusted testimony run as follows: 'That feeling is not really pain at all. Real pain is always due to the firing of C-fibres, but this alleged mild pain that you are feeling is not real pain at all but what we now call quasi-pain, which we have recently discovered is due to the firing of D-fibres'. In such a scenario, the subject could still retain the conceptual repertoire to form the belief that they are in pain and yet the given testimony is strong enough to defeat their justification—after all, they are not even very sure that they feel in pain because the feeling is only very faint. This line of thought undermines KL and WL, and indeed, if it is dialectically illegitimate to deny that if C obtains then C determinately obtains, then it undermines KDL and WDL too.

That leaves KCL and WCL as the last hope for a broadly Cartesian conception of the mental. Now suppose that the subject intensely feels pain and yet simultaneously fully trusts some allegedly expert testimony to the effect that they are not really in pain at all. In such a case, the subject still *involuntarily* forms the belief that they are in pain. However, it's hard to see how this belief could be justified in any robust sense because they fully trust the testimony (and indeed remain epistemically responsible in accepting the testimony given that they trust the speaker). Let us also assume that the subject fully recognises that: if they are in pain, then the 'expert' is mistaken. They believe the antecedent of this known conditional but not the consequent—indeed they believe the consequent to be false. Hence, their belief that they are in pain is highly irrational since the subject knows that this belief conflicts with other testimonial beliefs that they hold (though it is of course excusable to hold the belief that they are in pain given the involuntary nature of such a belief). On any plausible account of knowledge, warrant, and justification, this irrational belief cannot be warranted or cannot constitute knowledge. There remains, of course, *some* justification for the subject to believe that they feel pain, but this pro tanto justification is not enough for the subject to justifiably believe that they are in pain, that is, there is no sufficient justification for them to believe that they are in pain. Consequently, Conee's argument remains intact.

Is there another way of saving the Cartesian Conception?

## 12. *Weatherson on luminosity.*

---

<sup>29</sup> Steup (2009) makes just this kind of response to Conee's argument.

Weatherson (2004) endeavours to rescue the Cartesian Conception via a theory of self-knowledge which has the following four ingredients:

- (i) There is a distinction between two sorts of belief that one can form with respect to (core) phenomenal states: ‘judgmental beliefs’, which are formed via ‘inner perceptions’ (or inner awareness of the phenomenal state) and ‘phenomenal beliefs’, which are such that the phenomenal belief that one feels cold is simply (token) identical with the phenomenal state of one’s feeling cold.
- (ii) Knowledge requires safety. However, we must distinguish between:

(Content-Safety) For all cases  $\alpha$ ,  $\beta$ , a belief  $B$  with content  $p$  is safe in  $\alpha$  if and only if  $p$  is true in  $\beta$  (where  $\alpha$  is close to  $\beta$ ).

(Belief-Safety) For all cases  $\alpha$ ,  $\beta$ , a belief  $B$  with content  $p$  is safe in  $\alpha$  if and only if  $B$  is true in  $\beta$  (where  $\alpha$  is close to  $\beta$ ).

- (iii) Content Essentialism is false and so the content possessed by the belief  $B$  in a case  $\alpha$  need not be the same as the content possessed by  $B$  in a close case  $\beta$ .
- (iv) There is a sharp cut-off in the phenomenal transition between feeling cold and not feeling cold.

Weatherson grants that Williamson has shown that ‘if beliefs about a [phenomenal] state must satisfy content-safety then that state is not luminous’ (2004). However, he argues that belief-safety is better motivated than content-safety for the reason that belief-safety can explain why a true mathematical belief formed by a lucky guess cannot constitute knowledge, while content-safety cannot. In more detail, suppose one forms the (true) belief that  $128 + 235 = 363$  via a lucky guess. Since the content of this belief cannot be false then *a fortiori* it cannot be false in a close case and so the belief (trivially) counts as content-safe. However, given the denial of Content Essentialism, the content of this belief could easily have been different—for example, it could easily have been the false content that  $128 + 235 = 365$ . Thus, the belief is not belief-safe and so cannot constitute knowledge, given that knowledge requires safety.

With respect to the phenomenal case, suppose one forms the phenomenal belief  $B$  at time  $t$  with the content *that one feels cold*, but at time  $t+1$  second one does not feel cold. The belief  $B$  is not content-safe since in a close case it is false that one feels cold. Nonetheless,  $B$  is belief-safe since at time  $t+1$  either  $B$  ‘no longer exists, or now has the true content that [one] does not feel cold’ (2004).

Even if belief-safety is the right safety condition on knowledge, this falls short of securing the luminosity of our core mental states since belief-safety is not sufficient for knowledge. On this score, Weatherson maintains that it is ‘rather plausible that [ $B$ ] is knowledge, for [the subject] could not have had *this* belief without its being true’ (2004). More generally, given the nature of

phenomenal beliefs, as conceived of by Weatherson, such beliefs satisfy not only the following principle of infallibility:

(IN) For all cases  $\alpha$ , if in  $\alpha$   $s$  believes that  $C$  obtains then  $C$  obtains in  $\alpha$ ,

but also the principle of credulity CR. Ordinary judgmental beliefs about core phenomenal states, in contrast, do not satisfy CR, and may not even satisfy IN (see below).

There are at least three worries with Weatherson's response. The first worry is that Williamson does not in fact invoke content-safety in order to undermine the luminosity principle KL. Content-safety, plus the thesis that knowledge entails belief, entails, but is not entailed by, the KM-like thesis: for all cases  $\alpha$ ,  $\beta$ , if in  $\alpha$  one knows that  $p$  then  $p$  obtains in  $\beta$ , where  $\alpha$  is close to  $\beta$ . (This KM-like thesis is derivable from KM given that knowing entails being in a position to know.) Moreover, this KM-like thesis straightforwardly entails the KS-like principle for all cases  $\alpha$ ,  $\beta$ , if in  $\alpha$  one knows that  $p$  then in  $\beta$  it is not the case that: one believes that  $p$  and not- $p$ . However, the converse entailment needs something like principle B. So, to deny that content-safety is not a necessary condition on knowledge (on the grounds that our phenomenal beliefs about our core mental states are not content-safe but nonetheless constitute knowledge) should not bring Williamson's safety condition KS on knowledge into question at all. Rather, it should bring into question the employment of the doxastic limited discrimination principle B. Indeed, given the way that phenomenal beliefs are characterised by Weatherson it is no surprise that B should fail since these beliefs are such that they need not *spill over*, as it were, to a close (but distinct) case. In particular, across the sharp transition from feeling cold to not feeling cold, the phenomenal belief that one feels cold simply cannot be formed in the case where one does not feel cold. Hence, the introduction of belief-safety (and the attendant commitment to a denial of Content Essentialism) as a replacement for content-safety (in the case of phenomenal beliefs at least) is a red-herring.<sup>30</sup>

A second, and much more serious, worry is that, given Conee's anti-luminosity argument, it is far too quick to assume that a belief which satisfies both CR and IN has thereby sufficient pedigree to count as knowledge. At best, Weatherson has shown that our mental states are sufficiently distinctive to furnish us with a cognitive home because such states satisfy both CR and IN, but that falls well short of salvaging the Cartesian Conception (see below).

The final worry is more serious still. Weatherson concedes that one's judgmental belief that one feels cold and one's phenomenal belief that one feels cold can come apart. Indeed, he also concedes

---

<sup>30</sup> Weatherson's point that content-safety cannot account for why a mathematical true belief that  $p$  which is gained via guesswork cannot constitute knowledge can be extended to Williamson's safety principle KS (see Sainsbury 1997). It may well be that KS has to be replaced by Belief-Safety to make proper sense of these cases—but this in an independent point against KS.

that at a single time a subject could phenomenally believe that they feel cold and judgmentally believe that they do not feel cold. But then the state of feeling cold is not luminous since the countervailing judgmental belief acts as a defeater to knowledge that one feels cold.<sup>31</sup> To this worry, Weatherson says:

If [a subject] has the judgmental belief that he is not cold [together with the phenomenal belief that he is cold] then the phenomenal belief that he is not cold may not be knowledge, for it is plausible that the existence of a contrary belief defeats a particular belief's claim to knowledge. But that does not mean that he is not in a position to know that he is cold (2004, p.?).

This isn't right. The countervailing judgmental belief defeats the subject's knowledge and also simultaneously defeats the subject's epistemic state of being in a position to know. One is in a position to know (via method M) that C obtains only if one can readily come to know that C obtains by forming the belief (via method M) that C obtains, (where the formation (via M) of this belief does not itself generate any defeaters which would defeat one's evidence that C obtains). If one is in a position to know that one feels cold then one feels cold (given that being in a position to know is factive). But if one feels cold then one phenomenally believes that one feels cold according to Weatherson. Furthermore, there is no question of the formation of the phenomenal belief itself generating any defeat. So, when one is in a position to know that one feels cold (on the basis of being in the state that one feels cold) then one thereby automatically knows that one feels cold (on the same basis). So, if Weatherson concedes that the existence of a countervailing judgmental belief acts as a defeater to knowledge that one feels cold, he should also concede that this judgmental belief acts as a defeater to being in a position to know that one feels cold. Since this predicament is true of all our core mental states then Weatherson has not shown that such states are luminous.

### *13. Central credulity, central infallibility, and infallibility.*

If we are to take Conee's defeater anti-luminosity argument seriously then the project of locating an *epistemic* feature of our core mental states which is sufficiently strong to furnish us with a cognitive home is surely doomed. As we have just seen, Weatherson can retreat to the position whereby our core mental states are such as to satisfy both CR and IN—at least for Weatherson's 'phenomenal beliefs'. That renders such states sufficiently distinctive to provide a cognitive home. While such a gambit is arguably along the right lines, we should be sceptical of the distinction between judgmental and phenomenal beliefs because this distinction was introduced to secure the luminosity

---

<sup>31</sup> Indeed, a non-actual but easily possible countervailing judgmental belief is also sufficient to defeat knowledge.

of our core mental states. As we have just seen, such states are not luminous given the possibility of countervailing judgmental beliefs, and so the motivation for the distinction lapses.

A more promising strategy is to recognise only one kind of belief concerning our (core) phenomenal states—judgmental beliefs—but suitably weaken CR yielding the following principle of *central credulity*:

(CCR) For all cases  $\alpha$ , if C centrally (i.e. *intensely*) obtains in  $\alpha$  under conditions N then in  $\alpha$   $s$  believes (via M) that C obtains.<sup>32</sup>

Here method M can be broadly seen as introspection.<sup>33</sup> Conditions N include that (i)  $s$  is functioning normally (and so is not tired, drunk, delusional, and so on), (ii)  $s$  has the requisite conceptual capacities to believe that C obtains, and (iii)  $s$  has had sufficient time to wonder whether C obtains with the aim of forming the belief as to whether or not C obtains. Furthermore, CCR should be taken to be valid for both positive and negative core mental states.

A key premise in the quick argument from §3 is that a subject may fail to form a belief (via introspection under N-conditions) that they feel cold/do not feel cold when they do/do not feel cold midway across a phenomenal transition from feeling cold to not feeling cold. Given CCR, all that follows is that the subject neither intensely feels cold nor intensely feels not cold. When one intensely feels cold/not cold then (given that N-conditions obtain) then there is surely no scope for a subject to fail to form a belief (via introspection) that they feel cold/not cold. Indeed, were a subject to fail to form a belief that they feel cold when they feel intensely cold we should conclude that the subject is not functioning normally or does not have the requisite concepts or indeed is not using mere introspection as their method of belief formation (see below). For these reasons, CCR is by far the most plausible basis for a cognitive home we have considered so far.

It is also plausible that:

(R) If in  $\alpha$   $s$  believes (via introspection under conditions N) that C obtains then in  $\alpha$   $s$  does not believe (via introspection under conditions N) that C does not obtain.

Given R, and CCR, we then have: if in  $\alpha$   $s$  feels intensely cold (under N conditions) then in  $\alpha$  they do not believe (via introspection) that they do not feel cold. Likewise, if in  $\alpha$   $s$  does not feel intensely cold (under N conditions) then in  $\alpha$  they do not believe (via introspection) that they feel

---

<sup>32</sup> CCR seems to be more or less equivalent to Shoemaker's (1996, p. 52) weak-thesis of self-intimation, whereby 'it is the essence of a [core mental state]  $M$  that its presence or absence normally intimates itself to the person', at least if *normal* cases of instantiation or non-instantiation of  $M$  are here taken to be *central* (i.e. intense) cases of instantiation or non-instantiation.

<sup>33</sup> We can remain neutral as to whether or not introspection is perception-like. So, CCR remains plausible even if the state of feeling pain and the state of believing that one feels pain are distinct essences that do not bear any constitutive connection.

cold. If we contrapose these claims and state them in schematic form then we have the following principle of *central infallibility*:

(CIN) For all cases  $\alpha$ , if in  $\alpha$   $s$  believes (via  $M$  under conditions  $N$ ) that  $C$  obtains then it is not the case that  $C$  centrally fails to obtain,

(where CIN is valid for both positive and negative core mental states).

So, CCR entails CIN (given R), but not vice versa. Also, given that if  $C$  centrally obtains then  $C$  obtains, then IN, suitably indexed to a method  $M$  and to  $N$ -conditions, as follows:

(IN) For all cases  $\alpha$ , if in  $\alpha$   $s$  believes (via  $M$  under conditions  $N$ ) that  $C$  obtains then  $C$  obtains in  $\alpha$ ,

entails CIN, but not vice versa. So, if we can find a counterexample to CIN then both CCR and IN are invalid also.<sup>34</sup>

#### 14. *The initiation counterexample.*

As it turns out, there is a well-known (purported) counterexample to IN that also appears to be a counterexample to CIN (and thus to CCR too given R):

[...] surely it is possible to be mistaken about our bodily sensations? Take the initiation trick where the unfortunate subject is blindfolded and told that he will be branded with the red hot poker he can hear, and smell, being heated for the purpose. A piece of ice is pushed into his bare stomach and, naturally enough, he screams. There seems a reasonable case for saying he mistakes a sensation of cold for one of warmth. To be sure he will probably realise the deception almost immediately, but if he dies of fright on the spot it seems he will die thinking that he feels a sensation of intense heat, or pain, when in fact he does not.<sup>35</sup>

This represents a putative counterexample to CIN (and so IN) because it seems that the subject believes that they feel a sensation of heat and yet they centrally feel a sensation of cold, and so, centrally do not feel a sensation of heat. There are at least four ways in which one might resist this counterexample:

(i) Jackson (1973 pp. 59-60) disputes that the subject forms the belief that they feel a sensation of heat because while the subject believes he is being touched by a hot poker, and while it is true

---

<sup>34</sup> Note that Conee's defeater anti-luminosity argument also serves to undermine the thesis that for all cases  $\alpha$ , if in  $\alpha$   $s$  believes (via  $M$  under conditions  $N$ ) that  $C$  obtains then in  $\alpha$   $s$  knows that  $C$  obtains (and so knows that it is not the case that  $C$  intensely fails to obtain).

<sup>35</sup> Don Locke (1967, p. 86), quoted from Jackson (1973, p. 59).

that red hot poker causes painful sensations of intense heat, it does not thereby follow that the subject believes that he (intensely) feels a sensation of heat. Here Jackson appeals to the obvious invalidity of the schema: If *s* believes *p*, and *p* entails *q*, then *s* believes *q*. This is an unconvincing reply. The subject also surely *believes* that red hot pokers cause painful sensations of intense heat. Given that they also believe that they are being touched by a red hot poker, then, on that basis they also surely believe that they intensely feel a sensation of heat.

(ii) A somewhat more promising reply is to hold that the subject does not in fact feel a sensation of cold (intense or otherwise) but instead (intensely) feels a sensation of heat and, moreover, that this sensation occurs in virtue of the fact that they (intensely) *believe* that they feel such a sensation. In other words, there can be certain circumstances under which the obtaining (or non-obtaining) of a sensation can be response-dependent. Part of the story here will be to recognise the effect that background beliefs and expectations have upon the phenomenal character of experience. Though this reply has some mileage in it there are better options available.

(iii) On a more concessive response, it can be conceded that the subject mistakes a sensation of heat for one of cold. However, this mistake is rectifiable given time. Indeed, in the example as set up by Don Locke above, it is conceded that the subject 'will probably realise the deception almost immediately'. One of the N-conditions is that the subject is given sufficient time to wonder whether they are in a core mental state. So, it can be allowed that a momentary mistake has been made but this merely shows that the N-conditions have not been fully met and so CIN/IN remain intact.

(iv) On a related reply, (the one that I favour), it may also be acknowledged that the subject believes that they feel a sensation of heat when in fact they (intensely) do not but that this only shows that the belief has been acquired on a non-standard basis. In particular, the basis for this belief is not introspection *simpliciter*, but introspection mediated (and indeed confounded) by a set of background beliefs which includes the misleading belief that one is about to be touched by a red hot poker. The thought is that in cases where the subject uses introspection *simpliciter*, no counterexample to either CIN or IN emerges. This is especially so, if introspection *simpliciter* is taken to provide some sort of direct, immediate access to our (core) mental states (a feature which is still compatible, of course, with a perceptual model of introspection). The upshot is that the initiation counterexample provides no reason to reject either CIN or IN.

Given the unpersuasiveness of the initiation counterexample, it might be thought that we should embrace IN in addition to CIN and thus have a stronger and so more distinctive conception of what it is to have a cognitive home. The credulity principle CR, together with principle R, entail the infallibility principle IN. However the converse entailment is not valid.<sup>36</sup> This then opens up the

---

<sup>36</sup> Shoemaker (1996, pp. 51-52, p. 69) claims that, given R, principle CR and IN are interderivable. Despite having sympathy for Shoemaker's endeavours to rehabilitate a limited form of Cartesianism, I find it odd that he casts some

possibility of retaining IN while rejecting CR.<sup>37</sup> However, there is a good reason to resist IN: even if CR is replaced with CCR (because of the possibility of ‘first-person agnosticism’), nonetheless, a fully competent, normally functioning, fully attentive subject may nonetheless still form beliefs about whether or not they are in a core mental state across the borderline area. Under the supposition that a non-epistemicist theory of vagueness is correct, such beliefs will be indeterminate, and so defective, in truth-value: IN will have a determinately true antecedent and a consequent which is not determinately true. On either a standard supervaluational or a standard three-valued conception of vagueness, IN will thus be invalid.<sup>38</sup> Under the supposition that an epistemic conception of vagueness is correct, whereby it is readily granted that introspection is less than perfectly discriminatory, such beliefs will sometimes turn out to be false in the borderline area. Either way, IN, but not CIN, is suspect.<sup>39</sup>

To rescue IN from the worry just raised it would be necessary to postulate some sort of constitutive connection between the belief that one is in a core mental state and the fact that one is in that state. Whether, and in what way, such a connection obtains is an issue which lies beyond the scope of the present paper. Moreover, as we have seen, to establish that we are cognitively homed there is no need to demonstrate that our core mental states satisfy IN. Satisfaction of CCR, and so CIN, is arguably sufficient. The upshot is a non-epistemological, purely cognitive conception of what it is to have a cognitive home. In particular, the issue as to whether or not we are cognitively homed has been divorced from whether or not our core mental states are luminous, cozy, centrally luminous, quasi-luminous, warrant-luminous, warrant-cozy, and so on. Moreover, Williamson himself need have no deep objection to the position being advanced because it is perfectly possible to reject the phenomenal conception of evidence and yet retain the thesis that we have a cognitive home in the sense argued for.<sup>40</sup>

---

doubt on IN and yet defends CR (and not the other way around) given that the latter principle entails the former but not vice versa.

<sup>37</sup> Chalmers (2003) defends a specific form of IN as applied to what he calls ‘direct phenomenal beliefs’. However, for Chalmers, beliefs about whether or not one feels pain, feels hot, and so on, do not count as direct phenomenal beliefs, and indeed Chalmers seems to accept that the initiation counterexample shows that IN fails for such non-direct phenomenal beliefs. Horgan and Kriegel (2007) likewise defend a limited form of IN under which, very roughly, a phenomenal belief that one’s experience *feels like this* cannot be mistaken (where ‘this’ does not function as a mere indexical but as a particular kind of phenomenal mode of presentation). Like Chalmers, they do not think that beliefs about whether or not one is in a core mental state are infallible (though in their footnote 17 they do express strong doubts about the initiation counterexample).

<sup>38</sup> On the most plausible many-valued conception, IN will take the truth-value taken by the consequent (given that the antecedent is true to degree 1). Hence, IN could feasibly come out as less than half true (and so effectively invalid).

<sup>39</sup> A further reason to resist IN concerns the Wittgensteinian complaint that ‘if you can’t be wrong, you can’t be right either’—hence IN is somehow a conceptually incoherent principle. This position is endorsed by Armstrong (1963, p. 422) and in a modified form by Wright (1989, p. 634). Note that this complaint does not affect the standing of CIN, for this principle can allow that beliefs about whether or not one is in a core mental state can be mistaken.

<sup>40</sup> Indeed it is no part of the view on offer than knowledge is a *core* mental state for one can centrally know that *p* without it being the case that one believes via introspection that one knows that *p*. So Williamson can retain his thesis that knowledge is a mental state and yet accept that we do have a cognitive home.

Even if one remains thoroughly suspicious of Conee's defeater argument, nonetheless Williamson's anti-luminosity argument, and the anti-cozy variant, give us reason to think that our core mental states are neither luminous nor cozy. If, in addition, one is suspicious of these arguments, there is still scope to argue (via WWM) that our core mental states fail to be warrant-luminous, and so fail to be luminous. Should one, in addition, doubt the persuasiveness of this further argument then the possibility of first-person agnosticism invoked in the quick argument provides an alternative (and arguably much better) reason to think that our core mental states are not warrant-luminous and so not luminous. Indeed, if we are to remain neutral as to the correct theory of vagueness (and in particular as to whether or not 'feels hot' draws a sharp boundary) then the quick argument also shows us that the only plausible epistemological conception of what it is to have a cognitive home resides in the claim that our core mental states are centrally luminous. Such a position falls well short of the Cartesian Conception as originally conceived. If, on the other hand, Conee's argument is persuasive (as was argued above) then the Cartesian Conception must be abandoned in both letter and spirit. It is much easier to do that, and in turn much easier to embrace Conee's argument, once it is realised that this does not entail abandoning the idea that each of us has a cognitive home.<sup>41</sup>

## REFERENCES

- Armstrong, D. (1963): 'Is Introspective Knowledge Incorrigible?', *Philosophical Review*, 62, pp. 417-432.
- Brueckner, A., and Fiocco, M. O. (2002): 'Williamson's Anti-luminosity Argument', *Philosophical Studies*, 110, pp. 285-293.
- Chalmers, D. (2003): 'The Content and Epistemology of Phenomenal Belief', in Q. Smith and A. Jokic (eds) *Consciousness: New Philosophical Perspectives*, Oxford and New York: Oxford University Press.
- Cohen, S. 'Luminosity, Reliability, and the Sorites', ms.
- Conee, Earl (2005): 'The Comforts of Home', *Philosophy and Phenomenological Research*, 70, pp. 444-451.
- Fine, K. (1975): 'Vagueness, Truth, and Logic', *Synthese*, 30, pp. 265-300.
- Greenough, P. (2003): 'Vagueness: A Minimal Theory', *Mind*, 112, pp. 235-181.

---

<sup>41</sup> Acknowledgments ...

- Greenough, P. (2005): 'Contextualism about Vagueness and Higher-Order Vagueness', *Proceedings of the Aristotelian Society*, Supplementary Volume, 78, pp. 167-190.
- Greenough, P. (2008): 'Indeterminate Truth', *Midwest Studies in Philosophy*, P. French (ed.) *Truth and Its Deformities*, forthcoming.
- Hawthorne, J. (2005): 'Knowledge and Evidence', *Philosophy and Phenomenological Research*, 70, pp. 452-458.
- Horgan, T. and Kriegel, U. (2007): 'Phenomenal Epistemology: What is Consciousness that we may Know It So Well', *Philosophical Issues*, 17, pp. 123-144.
- Jackson, F. (1973): 'Is there a Good Argument Against the Incorrigeability Thesis?', *Australasian Journal of Philosophy* 51, pp. 51-62.
- Keefe (2000): *Theories of Vagueness*, Cambridge: Cambridge University Press.
- McGee V. and McLaughlin, B. (1995): 'Distinctions without a Difference', *Southern Journal of Philosophy*, 33, (Supplement), pp. 204-51.
- Neta, R. and Rohrbaugh, G. (2004): 'Luminosity and the Safety of Knowledge', *Pacific Philosophical Quarterly*, 85, pp. 396-406.
- Raffman, D. (1994): 'Vagueness Without Paradox', *Philosophical Review*, 103, pp. 41-74.
- Sainsbury, M. (1997): 'Easy Possibilities', *Philosophy and Phenomenological Research*, 57, pp. 907-919.
- Shapiro, S. (2006): *Vagueness in Context*, Oxford: Oxford University Press.
- Shoemaker, S. (1996): *The First-Person Perspective and Other Essays*, Cambridge: Cambridge University Press.
- Sosa, E. (2009): 'Timothy Williamson's *Knowledge and Its Limits*', in P. Greenough and D. Pritchard (eds) *Williamson on Knowledge*, Oxford: Oxford University Press.
- Steup, M. (2009): 'Are Mental States Luminous?', in P. Greenough and D. Pritchard (eds) *Williamson on Knowledge*, Oxford: Oxford University Press.
- Weatherson, B. (2004): 'Luminous Margins', *Australasian Journal of Philosophy*, 83, pp. 373-383.
- Williamson, T. (1994): *Vagueness*, London: Routledge.
- Williamson, T (1995): 'Definiteness and Knowability', *Southern Journal of Philosophy*, 33, (supplement), pp. 171-191.
- Williamson, T (1996): 'Cognitive Homelessness', *Journal of Philosophy*, 93, pp. 554-573.
- Williamson (2000): *Knowledge and Its Limits*, New York: Oxford University Press.
- Williamson (2005): 'Replies to Commentators', *Philosophy and Phenomenological Research*, 70, pp. 469-491.
- Wright, C. (1989): 'Wittgenstein's Later Philosophy of Mind: Sensation, Privacy, and Intention', *Journal of Philosophy*, 86, pp. 622-634.

