



How to Know That You're Not a Zombie

Brentyn J. Ramm¹

Received: 20 December 2022 / Accepted: 7 December 2023
© The Author(s) 2024

Abstract

I am aware of the tree and its leaves, but am I aware of my awareness of these things? When I try to introspect my awareness, I just find myself attending to objects and their properties. This observation is known as the ‘transparency of experience’. On the other hand, I seem to directly know that I am aware. Given the first observation, it is not clear how I know that I am aware. Fred Dretske thought that the problem was so acute that he issued the challenge of answering ‘How do you know that you are not a zombie?’ I propose that a view found in the Advaita Vedanta, that awareness is self-luminous, reconciles these two observations. I understand self-luminosity as the thesis that: (1) I am implicitly aware of my awareness and (2) I am phenomenally aware of a distinct phenomenal character of my awareness. In support of the first claim, that I apparently only attend to objects in the world when introspecting perceptual experience, suggests that I do not know my awareness explicitly, but rather that I must know it implicitly. In support of the second claim, I argue that the mere fact that I am perceptually conscious is not sufficient to allow me to know that I am perceptually conscious. In particular, the qualities I am perceptually aware of do not tell me that I am aware of them, rather they just seem to be properties of objects. I also assess whether strategies for responding to Cartesian sceptical scenarios can be employed against Dretske’s consciousness scepticism. I argue that these strategies either fail to distinguish me from a zombie or they do not adequately describe my epistemic situation. By contrast to other accounts, if awareness has its own distinct phenomenal character, then it cannot be considered to be a *prima facie* property of the world, hence the self-luminosity of awareness provides a plausible account of how I know that I’m not a zombie.

✉ Brentyn J. Ramm
Brentyn.Ramm@uni-wh.de

¹ Department of Psychology and Psychotherapy, Witten/Herdecke University,
Alfred-Herrhausen-Str. 50, 58448 Witten, Deutschland

1 Introduction

I am aware of the tree with its mottled, brown bark and the vivid greens of its leaves. I am also aware of my body, thoughts and feelings. These are all objects of awareness (in a broad sense of object). But am I aware of my awareness of these things? If I am, I do not seem to be aware of it in the same way that I am aware of the tree and its leaves. I cannot attend to it as I do for objects of awareness.¹ As G. E. Moore (1903, p. 450) famously put it, my awareness is ‘diaphanous’. I look for it, but I only find objects of awareness. In this same well-known section, Moore goes on to say that awareness itself can nevertheless be recognised – ‘it can be distinguished if we look attentively enough, and if we know that there is something to look for’ (1903, p. 450). Yet Moore did not give any further hint as to how he knew it. Fred Dretske thought that the problem is so acute that he issued the challenge of ‘How do you know that you are not a zombie?’ (Dretske, 2003a). The solution that I will propose here is inspired by the claim of some contemplative traditions that agree with Dretske that you cannot attend to your awareness, but hold rather that awareness is self-luminous (Albahari, 2009; Deikman, 1996; Gupta, 1998; Ho, 2007; Mackenzie, 2007; Thompson, 2014). There is an implicit background sense that you are aware — in fact, it is built-in to every conscious episode. My claim is that this provides an explanation of how you know your awareness. It answers Dretske’s question, as well as explaining why awareness is seemingly so elusive as Moore observed.

A phenomenal zombie is typically defined in the literature as a being that is physically and functionally identical to myself (or some other conscious being) but lacking in phenomenal consciousness (Chalmers, 1996). The notion that zombies are conceivable or indeed possible at all is controversial. Dretske, however, uses the term ‘zombie’ in a broader sense to refer to a human-like non-conscious being, while being neutral on whether such a being is (or could be) physically and functionally identical (functionalists tend to say it’s not possible, and Dretske also implies this by referring to the being as ‘human-like’). In fact, for Dretske a zombie is just an entertaining rhetorical device for referring to the problem of how you know you’re in a conscious state rather than a non-conscious state (Dretske, 2003a, note 1, p. 9–10). ‘How do you know you are not a zombie?’ can hence be used interchangeably with ‘how do you know you are conscious?’ The arguments below can hence be restated without mentioning zombies at all. I will use zombie here in the same broad sense as Dretske.²

The notion that you may not know that you are not a zombie (i.e., don’t know that you are conscious) may seem absurd, yet how you know this is not entirely clear.

¹ I will use ‘attend’ as a shorthand for an act of voluntary focal attention, to keep open the possibility that every aspect of conscious experience is at least partly attended, that is that background consciousness is in the periphery of attention. See Sect. 2.

² In reply to the objection that even Dretskean zombies are impossible, the question is why think this? With the recent success of large language models in mimicking human conversational ability, the possibility of Dretskean zombies seems to be even more plausible than ever. Furthermore, to the extent that there is a problem of how you know you are conscious, there is already some motivation for thinking that Dretskean zombies are possible, independently of one’s theoretical commitments about consciousness.

Dretske (2003a) grants that you know that you are aware of a tree and its properties, but asks how do you know that you are aware of it?

What makes us so different from zombies are not the things (objects, facts, properties) we are aware of but our awareness of them; but this, our awareness of things, is not something we are, at least in perceptual experience, aware of. So if you are, as you surely are, aware that you are not a zombie – aware, that is, that you are aware of things – what is it you are aware of that tells you this? Dretske (2003a, p. 2).

On the face of it in perceptual experience, when I try to introspect my experience, I focally attend to properties of the world, not properties of consciousness. This observation is known as the transparency of consciousness. I seemingly ‘look through’ my perceptual awareness directly to the world (e.g., Harman, 1990; Tye, 1995). But I also seem to directly know that I am aware. I will argue that the self-luminosity of awareness reconciles these two observations and hence provides a response to how I know my awareness.

It should be acknowledged from the outset that to what extent consciousness is ‘transparent’ and what this means is controversial in the philosophy of mind (see Kind, 2003; Siewert, 2004; Stoljar, 2004; Thompson, 2007, p. 282–287). I argue in Sect. 4 that Dretske’s challenge is weakened, but not adequately addressed by appealing to potential exceptions to transparency and the possibility of a weaker form of transparency in which it is difficult but not impossible to focally attend to my experience (Kind, 2003). One can of course appeal to other forms of consciousness such as thoughts and emotions, for which transparency is less compelling, but then this doesn’t explain the intuitive sense that at all moments you know in some sense that you are perceptually aware. I do not need to focally attend to my thoughts or emotions to know that I am aware—at least that’s the claim.³ I hence restrict the present arguments to knowing perceptual consciousness.

The arguments presented here are also somewhat unusual in that they attempt to move from transparency to a more complex account of consciousness, rather than the more usual direction of trying to reduce it to physical properties or relations (e.g., Dretske, 1995; Kennedy, 2009; Lycan, 2001; Tye, 1995). I also admit to being drawn to the flat-footed response to Dretske that the qualitiveness of experience is just self-evident. There something it is like for me to see the tree and nothing it is like for the zombie. I agree that knowing that you are conscious is a Moorean fact (e.g., Smithies, 2019, p. 159), but we can still ask how you know it. I think that the kernel of truth to Dretske’s (2003a) often mind-boggling argument is that it potentially links up with a view in found in Asian philosophy (and G. E. Moore) that redness, pain, afterimages etc. are mere objects of awareness, not awareness itself (Fasching, 2008). In this sense, the fact that there are objects and qualities being presented does not by itself tell me that *I am aware* of those objects and their qualities (at least not directly). My interpretation of Dretske’s challenge can be phrased as:

³ For arguments that non-perceptual experiences such as emotions, imagination and inner speech are also transparent, see Bryne (2018).

My awareness of perceptual qualities is necessary but not sufficient to put me in a position to know that I am perceptually aware.⁴

My claim is that I can focally attend to the particular qualities of my experience, but I only know I am aware of them because there is also a simultaneous implicit awareness of the awareness of that quality. Both the qualities and the self-luminosity of awareness are required to know that I am perceptually conscious. Another way of stating my claim is that I need to be aware that the properties are being *presented to me* to know that it is a conscious experience. Dan Zahavi refers to this dimension of conscious experience as the ‘first-personal givenness’ of experience (Zahavi, 2005, chapter 5).

To understand the notion of awareness being used here, as already mentioned, we need to distinguish between objects of awareness and ‘awareness itself’. In investigating consciousness, philosophers are usually concerned with explaining ‘qualia’, that is the qualitative properties of conscious experience. Examples include what it’s like to smell freshly cut grass, hear a plane drone overhead and feel the sharp pain of stubbing one’s toe. However, as recognised by G. E. Moore (1903) a phenomenal state has at least two components: (1) The objects and properties of which one is aware and (2) the consciousness of these objects and properties. He says ‘When we try to introspect the sensation of blue, all we can see is the blue: the other element is as if it were diaphanous. Yet it can be distinguished if we look attentively enough, and if we know that there is something to look for’ (p. 450).⁵ By ‘consciousness’, Moore did not mean the qualities of experience, but rather ‘the common element’ to all sensations (Moore, 1903, p. 450). This common element which he refers to as consciousness is *distinct* from qualities and objects of consciousness. It is elusive and diaphanous but it *can* nevertheless be recognised ‘if we look attentively enough’ (Moore, 1903, p. 450). If consciousness is the common element to all experiences as Moore suggests, then this seems to have the implication that ‘consciousness’ cannot be blue or green, loud or soft. Rather, to be compatible with all experiential qualities it needs to be neutral between them (e.g., Albahari, 2009). Hence, there is a potentially interesting resonance between Moore’s observations and claims from contemplative traditions that ‘consciousness itself’ is qualityless, that is, void-like, colourless and silent (Forman, 1999; Shear & Jevning, 1999). Furthermore, if awareness itself makes a distinct (i.e., non-sensory, non-affective and non-cognitive) contribution to conscious experience, then this is another means of resisting the claim of Dretske and others that perceptual awareness is exhausted by the presentation of worldly properties.

While awareness itself (or pure awareness) is most commonly thought to be recognisable in deep states of meditation (Metzinger, 2020) it is implausible that one

⁴ I am drawing the terminology ‘be in a position to know’ from Declan Smithies (2012, 2019).

⁵ Another interesting point of Moore’s quote is his claim that the distinction between objects of consciousness and consciousness itself can be recognised if we ‘look attentively enough’. Whether this means that Moore thought that awareness can be focally attended (in contradiction to transparency and the claim of the present article) or whether a special non-ordinary form of attention is required, is another question.

needs to meditate to know that one is aware. Rather that I am aware seems to be self-evident at every moment. This suggests that the sense that I am aware is an aspect of every experience. If awareness (understood as distinct from sensory, affective and cognitive properties) is indeed the essence of consciousness, then it must be present with every experience, whether I recognise it or not. Hence, even though I may overlook it, it seems that it must at least be implicitly present at every moment. How then do I know that I am aware?

The argument I will present here is epistemological. Here I focus on the claim that awareness is implicitly or pre-reflectively aware of itself. In Asian philosophical traditions this claim is put in terms of awareness being self-luminous (Gupta, 1998; Ho, 2007; MacKenzie, 2007; Thompson, 2014). The idea here is that awareness illuminates not just the qualities of the world, but also simultaneously illuminates itself. In the Buddhist tradition, awareness is sometimes compared to a lamp that lights up objects. At the same time, the lamp lights up itself. You don't need a second lamp to light up the lamp, rather it is inherently self-luminous (Ho, 2007). I claim that the self-luminosity of awareness provides an answer to Fred Dretske's sceptical problem of consciousness.⁶

The plan for the paper is as follows: In Sect. 2, I draw upon the notion of 'witness consciousness' in the Advaita Vedanta in fleshing out the self-luminosity thesis. In Sect. 3, I discuss the transparency of experience and how it supports implicit awareness of awareness (argument 1). In Sect. 4, I discuss the zombie problem for knowledge of awareness and how it supports the thesis that I know my awareness phenomenally (argument 2). In Sect. 5, I consider simple phenomenal accounts of knowing awareness and argue that they fail to answer Dretske's challenge. In Sect. 6, I argue that non-phenomenal accounts also fail to distinguish me from a zombie or are inadequate for accounting for my epistemic situation. In Sect. 7, I briefly address the highly counter-intuitive claim that I don't know that I'm aware. This I argue leaves the self-luminosity of awareness as the most plausible account in addressing Dretske's challenge.

2 Knowing Awareness

In the tradition of the Advaita Vedanta, pure awareness also known as 'witness consciousness' is an impersonal self-illuminating awareness. It is not the person, but rather the true subject. Its presence is known non-conceptually. It is the basis of all knowing, but is never known as an object (Albahari, 2009). The key epistemological claims of the view are summarised by Gupta (1998):

⁶ As well as the Advaita Vedanta and Buddhism (see the essays in Siderits, Thompson & Zahavi, 2011), arguments for versions of the implicit awareness of awareness can be found in phenomenology (Zahavi, 2005) and more recent analytic philosophy (Kriegel, 2009; Kriegel & Williford, 2006; Montague, 2016; Zahavi & Kriegel, 2015). Although I draw my inspiration from Asian philosophy, the present arguments do not necessarily decide between these different accounts.

1. The witness consciousness, although the basis of all knowing, is different from the object known. It is implied in every act of knowing. It is the ultimate subject; it can never become an object of knowledge.
2. It is the pure element of awareness in all knowing. It is one, immutable, indivisible reality.
3. It shines by its own light; it is self-luminous.
4. It is different from the empirical individual, who cognizes and enjoys. In other words, it is different from the empirical individual who is caught up in the triple states of waking, dreaming, and dreamless sleep. (Gupta, 1998, p. 18)

There are different possible forms of knowledge such as knowing-that (propositional knowledge), knowing-how, observational knowledge and innate knowledge. In the Advaita Vedanta, the self-luminosity of awareness is considered to be a non-conceptual form of knowledge. It is an ever-present and unmediated knowledge (Ram-Prasad, 2011, p. 234) and forms the basis for all knowing. If this is a form of knowledge (which would be controversial in the Analytic tradition), we could call it a form of knowing-as.⁷ There is no knower-known duality involved. A similar view is also found in the Buddhist tradition, particularly with the thinkers Dignaga and Dharmakirti (Siderits et al., 2011), the key difference being that the existence of the subject is denied. I think that knowing-as is a plausible form of knowing and I acknowledge it due to its importance in contemplative accounts of knowledge of awareness. However, here I will be predominantly concerned with conceptual knowledge, *knowing that* I am aware, as this allows us to link the self-luminosity account with Dretske's challenge. I understand the self-luminosity thesis as consisting of the following key claims in regards to conceptually knowing my awareness:

I am in a position to know that I am aware in virtue of the fact that:

- (1) I am implicitly aware of my awareness.
- (2) I am phenomenally aware of a distinct phenomenal character of my awareness.

This way of talking about our knowledge of consciousness draws on Declan Smithies' (2012, 2019) account of introspective knowledge with the difference that I phrase it in terms of awareness of awareness rather than qualitative consciousness. Like Smithies' view, this is a constitutive account of knowledge in which my justification for knowing my consciousness is infallible—whenever I am conscious I am justified in believing that I am conscious. I may fail to conceptually know that I am aware by lacking the appropriate concepts or if the correct concepts are not activated, hence the qualification, *I am in a position to know* that I am aware in virtue of the self-luminosity of awareness, rather than knowing it just in virtue of the awareness of awareness.⁸

⁷ For the case that there can be non-conceptual perceptual knowledge see Hoffman (2014).

⁸ See Smithies (2019, p. 158).

To understand the first claim of the self-luminosity thesis (above), we need to distinguish between explicit self-awareness and implicit self-awareness. One way to be explicitly aware of my conscious experience, such as my thoughts, is to think about them—thoughts are directed at other thoughts. The other aspect of explicit (or reflective) awareness is the act of attending. For example, when I am thinking about my thoughts, I also focally attend to them. They are in the foreground of my conscious experience. To be implicitly aware of something, by contrast, refers to an awareness of something in the absence of ordinary focal attention to it, and thoughts, feelings etc. about it. An example is the background hum of an air-conditioner. It makes a difference to my overall experience, even when I am not focally attending to it or thinking about it. A reason for thinking that my awareness is known implicitly is that if I could attend to it in the ordinary focal mode of attention, and bring it to the foreground of my consciousness, then it would be just another object of awareness before me, not awareness itself (Albahari, 2006, p. 6–8).

Some clarifications: By attention I refer to a form of selection, either conscious or unconscious. Another clarification to the above is that I refer to ‘focal attention’ rather than just attention, as it is controversial in the cognitive science literature as to whether or not all conscious experience requires attention. Some argue that there are degrees of attention and that every conscious experience, including outside of focal attention is at least partly attended, while others argue that consciousness and attention can occur without each other (for the debate see: Baars 1998; Block, 2014; Montemayor & Haladjian, 2015; Jennings 2015; Koch & Tsuchiya, 2007; Pitts et al., 2018). For those that hold that every conscious experience has some degree of attention, then the distinction between explicit and implicit partly rests on the former being whatever is in focal attention and the latter being whatever is in peripheral attention. I do not take a stand here as to which account is correct, but merely point out that there are two competing accounts of the relation between conscious experience and attention, both of which are compatible with the following arguments. There are also likely a number of varieties of explicit and implicit processes. For example, monitoring of our mental and attentional states as commonly occurs in meditation (meta-awareness) is arguably a non-propositional process which goes on in the background of consciousness (Dunne et al., 2019). This seems like a good candidate for an implicit process involving peripheral attention. In drawing the distinction between explicit and implicit in this way, I am assuming that there is conscious experience outside of the focus of attention (either unattended or peripherally attended or both). That is, I am assuming that the sense that there is a background to consciousness is not merely a kind of refrigerator light illusion in which the background aspects of consciousness are only there when I consciously ‘look for them’.

Another clarification is that I refer to an ‘ordinary’ mode of attention as I also think that there is a non-ordinary mode of attention (i.e., ‘inward attention’) in which awareness can be experienced explicitly but non-objectifyingly (Ramm, 2023; Shear & Jevning, 1999). For example, I previously argued that Douglas Harding provides such a means, by using a pointing finger, to assist one to attend inwards (in some sense of ‘attend’) and make salient a form of awareness that seemingly goes beyond sensory properties, i.e., a kind of ‘awake emptiness’ (Ramm, 2023). Independently of the truth of this claim, the proposal of the present paper is that you know that you

are aware even in the absence of such a non-standard explicit form of recognition, and that there must also be an implicit non-propositional form of recognition of this awareness (i.e., a self-luminous awareness).

The second claim of the self-luminosity thesis (above), as I understand it, is that the sense of being aware has its own distinct intrinsic phenomenal character (distinct from the qualities of objects of awareness). A pure awareness with no phenomenal character could not be known phenomenally and would be subjectively indistinguishable from no awareness at all (Albahari, 2009; Dainton, 2002; Ramm, 2023). The claim that awareness has its own distinct character can be found in Asian contemplative traditions such as the Upanishads which describes pure awareness as having a quality of luminosity (Thompson, 2014, p. 10) and Tibetan Buddhism in which it is claimed that awareness itself is a radiant or luminous emptiness (Padmasambhava, Coleman & Jinpa, 2006, p. 14–15).

3 Implicit Knowledge of Awareness

I will now move onto the argument for the first claim of the self-luminosity thesis, that I am implicitly aware of my awareness. The argument rests upon there being two plausible but seemingly incompatible theses about my everyday perceptual experience:

1. Apparent Direct Realism: In perceptual experience I am *prima facie* aware of objects in the world and their properties (not properties of awareness). When I see the tomato, I am *prima facie* aware of the redness of the tomato, not an inner redness *quale*.
2. 'I directly know that I am aware.

These are not metaphysical theses, but rather they are observations about our pre-theoretical experience. Whilst I think that there are good reasons for holding that direct realism is false as a theory of epistemology, it at least provides a good description of perceptual experience. When I try to attend to the green 'qualia' of my experience of the leaf, I find that I can only attend to the green property of the leaf, not apparently an inner subjective greenness.⁹ Whilst, on the face of it, I may not be able to directly attend to my awareness (only the objects of awareness), this leaves open the possibility that I know my awareness implicitly, that is outside of focal attention. This reasoning can be formulated into an argument for implicit awareness of awareness as follows:

⁹ This observation has been used to motivate the view that perceptual phenomenal character is identical to the represented properties of mind-independent things (Byrne, 2006; Dretske, 1995; Harman, 1990; Lycan, 2001; Tye, 1995, 2000) or mind-independent properties to which I am directly related (Campbell, 2002; Kennedy, 2009; Martin, 2002).

1. If in ordinary perceptual experience I know my awareness directly, then either (A) I am explicitly aware of my awareness or (B) I am implicitly aware of my awareness.
2. I am not explicitly aware of my awareness.
3. Therefore, if in ordinary perceptual experience I know my awareness directly, then I am implicitly aware of my awareness.
4. In ordinary perceptual experience I know my awareness directly.
5. Therefore, I am implicitly aware of my awareness.

Given the transparency of experience, I do not seem to be explicitly aware of my awareness. That is, I cannot attend to it as an object of my attention. This leaves only the possibility that I am aware of awareness implicitly, that is outside of focal attention. If the sense of being aware is a constant background presence, then I do not need to focally attend to it to know it. Knowing my awareness would not require introspection at all (roughly the act of focally attending to experience), and hence does not fall prey to the transparency observation. If my awareness is self-luminous, and hence implicitly known, it does not require voluntary, focused attention to know it. In fact, as previously mentioned, even if I could attend to my awareness as an object in front of me, it would merely seem to be an object of awareness, not awareness itself.

This argument relies on the premise that I am directly aware of my awareness. A challenge to the premise is accounts in which I know that I am aware indirectly. A prominent indirect account is that I know my conscious experience by inference. The strength of the inference explanation can be assessed by considering the other-illumination view, which is the main competing account to the self-illumination view. This view was put forth by in India by the Bhāṭṭa Mīmāṃsā epistemologists. According to this higher-order view, the objects that we are aware of have a property of cognizedness, such they are seen as available for speech, action and memory. On this basis there is an abductive inference that there is an awareness of the object (Siderits, 2011). Is this a good guide for inferring that there was an awareness of the object? There are at least four objections to this view.

Firstly, the classic objection, put forward by Dignāga is that the view leads to an infinite regress. If the result of the inference is an unconscious judgment, then it seems you cannot know it was made and hence do not know that you had an experience. If the resulting judgement is conscious, then it will itself require a further conscious judgment and so on. However, the other-illuminationist can deny that all cognitions, need themselves be cognised, and so the objection fails (See Siderits, 2011).

A second objection arises due to the time delay in the theory. Assuming that an inference takes some time and hence is not concurrent with the experience, this opens up the possibility that there was no awareness of the object at all. Perhaps, it is just a retrospective illusion that there was an experienced object (like the refrigerator light illusion). Hence it seems from this account that you cannot know that you were aware. Related to this is a phenomenological objection to the account, which says that surely, we seem to be aware of things, such as the blueness of the sky concurrently with it being presented to my awareness. If this is correct, then it seems that we need a concurrent awareness of awareness to account for this apparent

simultaneity. If the simultaneity is an illusion, then we are back to possibility that there you were not aware of the blueness of the sky, and hence do not know that you are not a zombie.

Thirdly, one can also question the validity of the abductive inference. As G. E. Moore points out ‘we can and must conceive that blue might exist and yet the sensation of blue not exist’ (1903, p. 445). This shows that from the fact that I detect the blue expanse of the sky, it does not follow that I am aware of it. Another way of putting it is that the inference does not distinguish me from a phenomenal zombie. Upon being unconsciously perceptually presented with the blue sky, the zombie may automatically infer that it is phenomenally aware of it. This inference is false in the case of the zombie. So, to infer from being presented with blue that you are aware of blue is a bad inference.

A fourth and perhaps most serious objection is that cognizedness does not seem to require consciousness at all. As an objective property which is out there, it is just a sheen on experienced objects, and hence could erroneously refer to a non-existent awareness.¹⁰ There is in fact neuroscientific and behavioural evidence that one type of cognizedness, availability for action, can be processed unconsciously. In one well-known study, it was found that when subjects are asked to grasp a disc that looks larger due to a size contrast illusion, subject’s grasp was not affected by the illusion; rather their grasp aperture adjusted to the object’s actual size (Agloti et al., 1995). There is also evidence from neuroscience that the dorsal pathway (which is specialised for visually guided actions) (Goodale & Milner, 1992) is activated when subjects are shown tools, even when subjects are unaware of the images (due to interocular suppression) (Fang & He, 2005). These findings suggest that this aspect of cognizedness can be processed and even guide some actions unconsciously. It’s not entirely clear what type of property cognizedness is, but one possibility put forward by Mark Siderits who provides a strong defence of the other-illumination view, is that it is a purely relational property. However, Siderits admits that on this interpretation there is no need for consciousness at all. Consciousness is just an empty place holder. Rather than taking this as a fatal implication of the view, Siderits embraces it and holds that it supports the reductive view that consciousness is a conceptual illusion. In other words, we are in fact zombies and the goal of Buddhist practice is to see through the illusion that consciousness exists. The other-illuminationists were certainly not reductionists about consciousness, so they would have resisted this conclusion. However, if Siderits’ analysis is correct, the other-illumination view does not allow us to know that we are not zombies. If so, this provides further support for the present argument that only on a self-illumination view can we directly know that we are conscious.

Another possible indirect route to knowing your experience is that you are aware of your experience via attending to the world (e.g., Dretske, 1995, p. 60–61; Tye, 2000, p. 52–53). As an analogy, I see that the fuel tank of the car is half full via attending to the fuel gauge. This would be an indirect account that does not use

¹⁰ Miri Albahari (2011) uses this metaphor to criticise analyses of for-me-ness which treat it as a mere reference within the stream of consciousness to a non-existing observing subject.

inference at all. Again, the problem is that this way of knowing your experience is also unreliable. Properties of the world can (on the face of it) exist whether or not they are being attended to. Zombies can also attend to properties of the world, but it would be false to say that they hence know that they are aware. This shows that there is no necessary connection between attending to properties of the world and knowing that you are aware. If it is true knowledge of awareness surely the connection is a necessary one. If I know anything it is that I am aware. Given the unreliability of inference and focal attention in knowing that you are aware, I believe that the indirect route doesn't adequately account for how we know our awareness. This supports the premise of the above argument that I know my awareness directly. As I do not seem to be aware of my awareness explicitly (e.g., by focally attending to it), this leaves only the implicit account. This argument hence provides support for the first component of the self-luminosity of awareness thesis, that I am implicitly aware of my awareness.

4 The Zombie Challenge and Phenomenal Knowledge of Awareness

I will now present an argument for the second component of the self-luminosity thesis that: I know my awareness as phenomenally distinct from the objects of awareness. The argument is as follows:

1. If I'm not phenomenally aware of a distinct phenomenal character of my awareness, then I don't know that I am not a phenomenal zombie.
2. I know that I'm not a phenomenal zombie.
3. Therefore I am phenomenally aware of a distinct phenomenal character of my awareness.

The aim of this section and Sects. 5 and 6 is to motivate the first premise. The aim of Sect. 7 is to motivate the second premise.

Dretske points out that I only seem to be perceptually aware of objects and their properties, not my awareness of them. Suppose I am looking at the blue expanse of the sea. I know that there is a blue expanse of water before me. According to one version of the transparency thesis, the water and its properties, a mind-independent material expanse, is all that I am aware of, not qualia and not my awareness. How then do I know that I am not a zombie? Since all I really know are objects of awareness, how do I know with certainty that I am actually phenomenally aware of them? If all that I perceptually register is objects and their properties, then on the face of it I have no reason to think that I am not a zombie. This directly conflicts with the other plausible thesis that I directly know that I am aware.

There have of course been exceptions proposed to perceptual transparency. Amy Kind (2003) distinguishes between Strong Transparency and Weak Transparency. Strong Transparency is the thesis that it is impossible to attend directly to your perceptual experience, while Weak Transparency is the thesis that it is difficult but not impossible to attend directly to one's perceptual experience. A

number of philosophers have proposed exceptions to transparency such as the phosphene-experience when pressing on one's closed eyeball (Block, 1996, p. 35), blurry vision (Pace, 2007) and hallucinations (Thompson, 2008). Perhaps when I have these experiences, I am aware of and can focally attend to my subjective experience rather than properties of the world. But surely, I don't need to press or squint my eyes or take LSD to know that I'm not a zombie. Accepting Weak Transparency as a position weakens Dretske's challenge but arguably doesn't adequately address it. It would mean that it is difficult but not impossible to know that you are not a zombie. In certain restricted circumstances, you can satisfactorily demonstrate to yourself that you are not a zombie. Surely though we know that we are perceptually conscious more easily than that. Some supporters of externalism will also deny that experiences such as hallucinations are subjective properties. Rather when you hallucinate you are related to uninstantiated mind-independent properties, not mind-dependent properties (Dretske, 2003b; Johnston, 2004; Tye, 2000). Thus, while the zombie does not experience hallucinations, it can presumably represent uninstantiated properties of the world, and hence can have proto-hallucinations and proto-illusions.

'But I have a perspective!' it may be objected. I can for instance see that a plate 'looks' elliptical when viewed at an oblique angle. Hence, Charles Siewert argues that I can attend to how things look to me and how they appear to me (Siewert, 2004). However, these 'modes of appearance' still only describe the way objects are manifesting to me, not the awareness itself. So how I know my awareness, arguably the essential component of a conscious experience, remains a mystery. Furthermore, your zombie twin also has a perspective of sorts. It tracks perceptually relative properties of the side of objects facing it. For example, because of the way light strikes its retina, its perceptual system registers something elliptical about seeing the plate, which it can also attend to. While the plate 'looks' elliptical to you, we could say that it 'proto-looks' elliptical for the zombie. Similarly, the zombie registers properties of the apple facing it but not its hidden side, and so it does not register the bite taken out of the hidden side. Again, in any case, presumably I don't need to view plates obliquely or to take special notice that things have unseen sides to know that I am aware.

But 'I feel pain!', one might insist. 'A zombie does not feel pain and so I am not a zombie.' Certainly, your awareness is lighting up the pain, and the painful qualities are all too evident. You are aware of painfulness (seemingly a property of your body), but are you aware of your awareness of the pain? Although my zombie twin does not feel pain it plausibly has proto-pain. There is a tracking of bodily damage, disorders and irritations. The zombie hence has proto-pains, proto-headaches and proto-tickles. You by contrast feel pains, feel headaches and feel tickles. You know this, but *how* do you know this?

In this section, I used transparency to call into question how you know your awareness. If on the other hand, I know my awareness itself phenomenally, then perceptual transparency (in its various degrees) is compatible with the direct knowledge of awareness. If awareness has its own unique phenomenal character, then it cannot be considered to be a *prima facie* property of the world. Hence, phenomenal knowledge of awareness provides a clear solution to Dretske's challenge.

5 The Simple Theory of Phenomenal Knowledge

Rather than looking for exceptions to transparency, another plausible way to deny that an awareness of a distinct sense of awareness is needed to know awareness is to appeal to conscious perceptual experience itself. The idea is that I while a zombie unconsciously represents objects, I by contrast consciously represent objects. I hence know that I am conscious via my conscious representation of things. A version of this alternative phenomenal account of knowing consciousness is Declan Smithies' Simple Theory of introspection (Smithies, 2012, 2019). According to this view, when I am aware of green, I have an infallible justification for believing that I am aware of green. This puts me in a position to know that I am aware of green just in virtue of being in that mental state (Smithies, 2012, p. 261). This provides an explanation for why one's knowledge of one's mental states is distinctive from perceptual knowledge. In particular, one's justification for believing one is presently having a particular conscious experience is infallible, in contrast to perceptual justification. As an example, I have justification for believing that there is a green leaf in front of me by perceiving it. This justification is not infallible because I could be hallucinating. That I seem to be seeing the greenness on the other hand is justified just in virtue of having that experience of green.

Importantly, this view allows that I phenomenally know that I am not a zombie, but without an implicit awareness of awareness with its own unique phenomenal character. In fact, awareness and phenomenal qualities are the same thing. Another way of putting it is that the difference between the simple view and the self-luminosity view is what gets 'lit up' in a conscious experience. According to the simple view only the qualities of consciousness are lit up. According to the self-luminosity view awareness lights up both the qualities of consciousness and itself.

I think that there is a lot that is right about the simple view. I agree that if you are aware then you have infallible justification for believing that you are aware and that this puts you in a position to know this. If we combine this view with a constitutive relation between the phenomenal belief and the experience (e.g., Chalmers, 2003; Gertler, 2001; Horgan & Kriegel, 2007), then I also have some degree of infallible conceptual knowledge that I am having experiences. However, whether or not the account helps respond to Dretske's challenge is going to depend upon whether it can distinguish between the following two cases:

The Good Case: I am aware of the tomato, its red colour and its bulbous shape.

The Bad Case: I unconsciously perceptually detect the tomato, its red colour and its bulbous shape.

If you know that you are in the good case then you know that you are not a zombie. You surely know that you're in the good case (Dretske grants this), but he asks how do you know it? Suppose that the difference between me and the zombie is that the colours represented in my perception have a special glowy-property

while the colours perceived by the zombie lack this glowy-property. The problem is that we seem to have access to the same type of information as a zombie. For example, we both have access to the apparently objective information that there is a tomato, with a certain shape and colour, at a certain distance. What the zombie and I have in common is that we both have access to the apparent properties of objects. On the face of it those glowy-colours also belong to the objects I see—they are for all I know objective properties. Perhaps the colours are actually metaphysically objective and hence exist without being perceived. However, I don't have to believe in colour objectivism to see that colours at least apparently belong to the objects I see. We are still talking about a property of things even if a very special property and as Dretske (2003a, p. 2, 9) points out an 'object of awareness' is not an 'awareness of an object'. So, what is it about those glowy-properties that tell you that you are aware of them? Not only do the properties not tell me that I am aware of them, they seem to be 'other' than me. The infallible justification provided by having an experience only helps you if you already know that you are in the good case. But how you know this is the very point of Dretske's question.¹¹

Daniel Stoljar (2012) denies that you need further evidence to know that you are perceptually conscious. Rather the perceptual consciousness itself provides all the evidence we need. He refers to this as the 'foot stamping response'. This response seems reasonable if you already know that you are not a zombie as Dretske indeed grants. If you know that you see the leaf, then you seeing the leaf is all the evidence you need. However, there are two forms of scepticism about knowing awareness. The weak form of scepticism, held by Dretske (2003a) concedes the point that you know that you are aware, but rather asks *how* you know it. If this reasoning is correct then 'the simple theory' and 'the foot stamping response' could provide elegant answers to Dretske (though Dretske of course disagreed). However, foot stamping doesn't address the transparency argument. If the properties I am aware of do not

¹¹ Michelle Montague (2016, chapter 4) puts the problem in terms of our natural ability to distinguish our experiences from the objects of experience. She similarly argues that without an implicit awareness of awareness we cannot make such a distinction, at least not with the ease we usually do. She draws upon an observation by P. F. Strawson that in everyday experience we naturally and unreflectively distinguish between our conscious perceptions of objects and the objects of perception. She suggests that any account that lacks an implicit awareness of awareness will struggle to account for this datum (particularly on a reading in which the distinction has its basis in phenomenology). Although I agree with her argument, it seems to me that opponents can more easily deny this datum than the fact that you know that you are conscious. Also, P. F. Strawson's datum that we naturally and 'unreflectively' distinguish between our experiences and the objects of experience seems to come a little too close to building the same-order awareness of awareness hypothesis into the datum. It is worth noting that my understanding of awareness of awareness is also different to that of Montague who characterises awareness of awareness as being aware of having a particular experience (such as a particular colour experience) just by having it (Montague, 2016, p. 108). This suggests there is a unique awareness of an experience for each particular qualitative property and object, so there will be numerous awarenesses of awareness occurring at any particular time. For me, by contrast, the sense of being aware is mode-neutral and singular, and encompasses the qualities both within and across sensory modalities at any particular moment all at once (such that there is only one self-luminous awareness at any particular moment as well as across moments).

seem subjective at all, and in fact could exist without any awareness at all, am I really justified in my belief that I am aware?

Foot stamping also doesn't provide a response to a stronger version of Dretske's challenge. If, based upon perceptual transparency, Dretske had not conceded that you know (somehow) that you are aware (strong scepticism) then this response would fail to answer the challenge. Afterall, if you don't already know whether you are in the bad case or the good case, then the justification you would have by being in the good case will not help you (Dretske (2003a) at times seems to be flirting with strong scepticism).

Finally, it could be objected that the self-luminosity thesis is just another version of the foot stamping response, but at a higher level. While this is true, what distinguishes the self-luminosity version of foot stamping from the simple account is that awareness itself is not a *prima facie* property of the world. It hence avoids Dretske's transparency argument. Another related objection is that a zombie could have proto-self-luminous states and also have as much conviction as I do that it is conscious. How then do I know that I'm not one of those zombies? Without the transparency motivation there is arguably little reason to believe that the self-luminosity of awareness is anything like that possessed by the zombie (I'm not in an analogous epistemic situation in this regard, as I arguably am for the qualities of perceptual consciousness). Unlike the zombie, according to the self-luminosity account, I know that I'm not a zombie because I am phenomenally aware of my awareness. Due to being acquainted with the unique phenomenal character of awareness, I am also justified in my belief that I am aware, unlike the zombie. This of course will not satisfy general sceptics about the existence of consciousness, such as illusionists, who deny that any phenomenality exists at all (see Sect. 7 below), however the goal here was the more modest one of answering Dretske's challenge.

6 Non-Phenomenal Knowledge of Awareness

By what non-phenomenal means could I know that I am aware? One possible non-phenomenal solution is that we know a state is conscious based upon its functional role. In fact, we typically at least partly distinguish between a conscious and an unconscious state by their functional differences. Conscious states are those that we can report, recall, and form beliefs and desires about. Unconscious states on the other hand tend to be states that we cannot report, recall or form beliefs/desires about. This distinguishes typical experiences from blindsight states where these functions tend to be missing. Blindsight is an ability in which subjects who are cortically blind can make prompted above chance guesses about stimuli presented to the blind field (e.g., whether there is an X or O, an object's location, colour etc.). This is the case, even though these subjects deny having any visual experience of the stimuli. Tye (1995) proposes that the difference between conscious perceptions and blindsight perceptions is that former but not the latter are functionally poised to affect the subject's cognitive systems, particularly their belief/desire systems. If the conscious/unconscious divide can be functionally defined as such, perhaps this will solve Dretske's challenge. You

cannot introspect whether your states are appropriately poised, but the outcome of these functions can be recognised (e.g., reportability). Perhaps then availability for report, categorisation, memory and beliefs/desires is how we know that we are conscious. ‘Knowing’ your conscious experience could even be defined in terms of cognitive accessibility.

There are two main problems with this proposal. The first is a conceptual problem. Ned Block (1995) makes a well-known distinction between phenomenal consciousness (what it is like to experience something) and access consciousness (availability for reasoning, speech and action) and argues that the two can come apart. He gives the case of being engaged in an intense conversation and suddenly becoming aware at noon that you have been hearing a loud drill outside for some time. The experience was ongoing in the background of consciousness but was not reportable or otherwise accessible until noon. We can also imagine varieties of super-blindsighters who either spontaneously make guesses about what is in the blind field or even just spontaneously form beliefs and desires about objects in their environment on the basis of unconscious perception (Block, 1995; Dennett, 1991, p. 343; Smithies, 2014). It seems then that the cognitive availability of a perceived object or property is not the same as being in a conscious state and is not sufficient for knowing that you are conscious.

The second and perhaps most serious problem is that cognitive availability seems to leave the original challenge unresolved. The functional difference does not solve the problem because Dretske is questioning how you know that you are conscious in cases where there is full cognitive access. The idea is that given transparency, you are just reporting, remembering and forming beliefs/desires about (apparently) objective properties of the world (not your subjective experience). As these properties can (apparently) exist whether or not you are aware of them, the question remains of how you know that you are (or were) aware of them. In other words, cognitive accessibility is beside the point. The epistemological question remains whether or not these states are cognitively accessible.

Finally, assume for the sake of argument that we know that Dretskean zombies and super-blindsighters are impossible (i.e., they must be conscious by definition). This would not solve the above epistemological problem because firstly the problem doesn’t rely upon this possibility (zombies are just tools for aiding one’s thinking about the problem). Secondly, deep philosophical thinking about zombies, blindsight and functionalism is surely not how you know that you are conscious. Pre-theoretically the problem would remain. If certainty that zombies and super blindsighters are not possible was the only way of knowing that you are conscious then any philosopher who does believe these entities are possible wouldn’t know they are conscious, and neither would most non-philosophers know that they are conscious!

Michael Tye (2009) has proposed another non-phenomenal solution to Dretske’s challenge. Tye (2009, p. 191–193) argues that Dretske’s consciousness scepticism is analogous to external world scepticism, and thus can be answered using similar strategies as for answering the Cartesian sceptic. Hence it is not a new epistemic problem. Cartesian external world scepticism is the challenge of how you know that there is an external, mind-independent world, rather than just a hallucinated or vividly dreamed world. Various responses to the Cartesian sceptic include: (1) direct

experience, (2) inference to the best explanation (Vogel, 1990), (3) no good reason to the contrary (Pryor, 2000) and (4) entitled belief (Wright, 2004).

To say that I know the world because I directly experience it is the equivalent of Dr Johnson kicking the rock in response to Berkeley's idealism and exclaiming 'I refute it thus!' (Campbell & Cassim, 2014, p. 25–26). This is not a particularly compelling response to either Berkeley or scepticism. I could kick a rock in the Matrix and it obviously would not be a proof that I am not in the Matrix. Still that I seem to experience the world, and given other's reports that they also experience it, it provides strong *prima facie* evidence that it continues to exist whether I am experiencing it or not. This however does not help when it comes to awareness. With the world there is an answer to how I know it, namely by perceptual experience. However, with awareness, on the face of it, my evidence is only about facts and properties of the world, not about awareness itself. Unless I am also aware of my awareness experientially it seems that this response to external world scepticism is not available for answering consciousness scepticism.

Another common response to external world scepticism is inference to the best explanation. According to this argument, the best explanation for the regularity of the world is that it continues in the form it seems to have when I am not perceiving it. Its laws continue in exactly the same fashion, such that my plant wilts when sitting in a hot room without water whether I am perceiving it or not. Bertrand Russell presents a version of this argument in *The Problems of Philosophy* (Russell, 1912). He points out that if the world doesn't continue to exist when I'm not looking at it, there is no explanation for why my cat gets hungry between meal times. This is an appeal to explanatory simplicity. However, if the only facts I am perceptually aware of are objects and properties of the world, either instantiated or uninstantiated, it seems that consciousness does not need to be introduced into the explanatory picture. It does not seem to be doing any explanatory work. It could be, as Siderits (2011) argues, an empty placeholder, that is just a conceptual illusion. The simplest and best hypothesis, if there is no awareness of awareness, seems to be that I am a zombie—there is no consciousness at all.

Finally, perhaps I have an entitled belief that I am aware. Suppose I am looking at a zebra in the Zoo. Do I know that it is a zebra? One would be tempted to say yes, but it is possible that it is actually a cleverly described mule, so I do not know with absolute certainty that it is a zebra (Dretske, 1970). However, given that I am in a zoo, and that conspiracy theories are highly complex, I at least have an entitled belief that it is a zebra. Perhaps this is the way I know my awareness. James Pryor (2000) argues similarly that perceptual experience provides defeasible justification that the external world exists. In the absence of a good reason to the contrary, my perceptual experience provides justification for believing that the external world exists. Interestingly, on Pryor's account you do not need to be *aware* of your experiences to have this justification. Rather having an experience as of an apple is enough to justify one's belief that there is an apple (Pryor, 2000, p. 519). Importantly, according to this 'dogmatist' view, no further non-question-begging reasons or evidence apart from just having the experience is needed to supply justification.

The first thing to note about applying the entitled belief and dogmatist responses to consciousness is that they only provide defeasible justification that I am aware.

The justification could be undermined by learning that it was mistaken. This however doesn't seem to be adequate in explaining my epistemic situation. If someone really presses me about whether I know that it is a zebra, I will concede that I do not know in an absolute sense of 'know'. I may also concede this with the 'external' world, especially directly after watching *The Matrix*. However, such doubts do not seem to be admissible for whether I know that I am aware. As Descartes observed: 'I am now seeing light, hearing a noise, feeling heat. These objects are unreal, for I am asleep; but at least I seem to see, to hear, to be warmed. This cannot be unreal' (Rene Descartes, 1971, Meditation Two, p. 71). Even if we disagree with everything else Descartes said, his conclusion here is rock solid—at least to my mind.

As already discussed, Strong Transparency is controversial, however, if the arguments presented here are right it suggests that *if* perceptual consciousness is completely exhausted by the presented properties of the world, then I don't know that I am perceptually aware. Even if transparency is only partly true, it suggests that it is more difficult to know that I'm conscious than we would intuitively expect. If on the other hand, awareness is self-luminous, such that there is a unique (i.e., non-sensory, non-affective and non-cognitive) phenomenal character to the sense of being aware, then there is an aspect of consciousness that goes beyond presented properties of the world. This would provide an explanation to how I know my consciousness and why my awareness is self-evident at every moment. In the next section, I briefly consider the attempt to block the argument by rejecting the second premise.

7 Denying that You Know that You're Not a Zombie

In the previous section, I assessed a number of non-phenomenal proposals for how I know that I am aware and found them wanting. I conclude that non-phenomenal means of knowing are not sufficient to know that I am not a zombie. In lieu of finding a non-phenomenal means of knowing awareness which I could possess but not my zombie twin, the alternative is bite the bullet and to deny premise 2, that is deny that I know that I am not a zombie. Of course, it may be the case that I am conscious, even if I cannot know this fact for certain. Thus like the external world sceptic the denier of premise 2 is not denying that they are conscious, rather they are withholding judgement. I take it that this would be a conclusion that most would like to avoid and I have generally assumed that it is obvious that I know that I am conscious, and so the question was how exactly I know this.

This being said a number of philosophers have argued for illusionism about consciousness which is the thesis that consciousness only seems to exist, but actually does not (Kammerer, 2018; for a volume dedicated to illusionism see Frankish, 2017). The attraction of illusionism for some materialists is that explaining our intuitions/judgments about why consciousness seems to exist is arguably more naturalistically tractable than explaining phenomenal consciousness within a materialist framework. Dretske's challenge, particularly the strong version, is another argument that can be wielded by the illusionist. If I really do not know that I am aware, then why posit that awareness exists in the first place? There is no access to my awareness from the third-person perspective, only my reports that I am aware. It is only

from a first-person perspective that I know that I am aware. If for all I know from my perspective I could in fact be a zombie, then there would be no reason left to even posit the existence of awareness.

I do not have space to adequately respond to illusionist arguments here. My own position is that the justification of our beliefs about the existence of the world are non-inferentially grounded in our conscious perceptual experiences. A zombie by contrast does not have justified beliefs about the world (for an extended argument see Smithies, 2014, 2019). As illusionist arguments rely heavily upon an appeal to science and if our judgments about the world (and hence science in general) are justified by conscious perceptions as Smithies argues, then one of the main motivations for the illusionist view would be self-undermining. The findings of science cannot undermine the existence of consciousness if those very findings depend upon conscious observations. Furthermore, like many I hold that if anything is self-evident it is that I am conscious. This is a version of the Moorean response to illusionism (Chalmers, 2020: for an illusionist response see Kammerer, 2022). I believe that Dretske would have agreed. Even if we take a Moorean position, the question still remains of how exactly we know that we are conscious. Another potentially significant outcome of the current arguments is that, if they work, they seem to imply that either the self-luminosity thesis is true or scepticism about knowledge of consciousness and/or illusionism about consciousness is true. The main challenge for illusionists is explaining why our intuitions that we are conscious are so strong, particularly the Cartesian intuition that it is undoubtable that I am conscious, and to find ever more creative ways of weakening these convictions.

8 Conclusion

I argued that the thesis that awareness is self-luminous provides an answer to Dretske's question of how you know that you're not a zombie. The problem arises because of a tension between the fact that (1) I seem to be aware of properties of external objects (rather than properties of perceptual experience), and (2) I seem to directly know that I am aware. If the sense of being aware has a distinct phenomenal character, then it is not a *prima facie* property of the world. Importantly, the thesis can resolve this tension independently of whether or not all properties of perceptual experience are world-presenting.

Many will insist that the greenness of the leaf that I am aware of is self-evidently qualitative and hence I am not a zombie. I am sympathetic with this reaction. Another way of saying it is that the greenness is a phenomenal property—end of story. To avoid confusion, my proposal is not just that awareness lights itself up, but that it simultaneously lights up the qualities of the world. These are inseparable aspects of conscious perceptual experience (see also Ho, 2007). However, presumably what makes the greenness of the leaf phenomenal isn't just its 'qualitativeness', but the fact that *there is awareness of the greenness*. If you take away the fact that there is awareness of the quality, then the 'phenomenal' greenness vanishes. There may be a mind-independent greenness out there somewhere, but it is not phenomenal green until I or some other subject am aware of it. Thinking about

zombies is helpful, though not necessary, for seeing the problem of how you know your awareness.

Experiences of ‘awareness itself’ have been reported in contemplative traditions across different times and cultures (Metzinger, 2020; Thompson, 2014), as well as by contemporary practitioners (Gamma & Metzinger, 2021). Many contemplative traditions distinguish between awareness and the objects of awareness. It is also common in these traditions to hold that awareness is self-luminous. Here I have endeavoured to provide an epistemological argument for this claim, and in particular to show that it provides a plausible answer to how I know that I’m not a zombie.

Acknowledgements Many thanks for helpful comments on a presentation of this paper from the attendees of the University of Western Australia Philosophy Society Seminar, particularly from Miri Albahari, Remco Heesen and Chris Letheby. Special thanks to David Pitt for helping me see how an early version of this argument could be made workable.

Funding Open Access funding enabled and organized by Projekt DEAL. This research was funded by a Postdoctoral Research Fellowship received by the author from the Alexander von Humboldt Foundation.

Declarations

Conflict of Interests The author declares no conflicts of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Aglioti, S., DeSouza, J. F., & Goodale, M. A. (1995). Size-contrast illusions deceive the eye but not the hand. *Current Biology*, 5(6), 679–685.
- Albahari, M. (2006). *Analytical Buddhism: The two-tiered Illusion of Self*. Palgrave Macmillan.
- Albahari, M. (2009). Witness-Consciousness: Its definition, appearance and reality. *Journal of Consciousness Studies*, 16(1), 62–84.
- Albahari, M. (2011). Nirvana and ownerless consciousness. In M. Siderits, E. Thompson, & D. Zahavi (Eds.), *Self, No Self? Perspectives from Analytical, Phenomenological, and Indian Traditions* (pp. 79–113). Oxford University Press.
- Block, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences*, 18(2), 227–247.
- Block, N. (1996). Mental paint and mental latex. *Philosophical Issues*, 7, 19–42.
- Byrne, A. (2006). Color and the mind–body problem. *Dialectica*, 60, 223–244.
- Byrne, A. (2018). *Transparency and Self-knowledge*. Oxford University Press.
- Campbell, J. (2002). *Reference and Consciousness*. Clarendon Press.
- Campbell, J., & Cassam, Q. (2014). *Berkeley’s Puzzle: What Experience Teaches Us*. Oxford University Press.
- Chalmers, D. J. (1996). *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press.

- Chalmers, D. (2003). The content and epistemology of phenomenal belief. In Q. Smith & A. Jolic (Eds.), *Consciousness: New philosophical Perspectives*. Oxford: Oxford University Press.
- Chalmers, D. (2020). Debunking arguments for illusionism about consciousness. *Journal of Consciousness Studies*, 27(5–6), 258–281.
- Dainton, B. (2002). The gaze of consciousness. *Journal of Consciousness Studies*, 9(2), 31–48.
- Deikman, A. J. (1996). I = awareness. *Journal of Consciousness Studies*, 3(4), 350–356.
- Dennett, D. C. (1991). *Consciousness Explained*. Little, Brown and Company.
- Descartes, R. (1971). *Descartes: Philosophical Writings*. Nelson's University Paperbacks.
- Dretske, F. (1970). Epistemic operators. *The Journal of Philosophy*, 67(24), 1007–1023.
- Dretske, F. (1995). *Naturalizing the Mind*. MIT Press.
- Dretske, F. (2003). How do you know that you are not a zombie? In B. Gertler (Ed.), *Privileged Access*. Ashgate Publishing Limited.
- Dretske, F. (2003b). Experience as representation. *Philosophical Issues*, 13(1), 67–82.
- Dunne, J. D., Thompson, E., & Schooler, J. (2019). Mindful meta-awareness: Sustained and non-propositional. *Current Opinion in Psychology*, 28, 307–311.
- Fang, F., & He, S. (2005). Cortical responses to invisible objects in the human dorsal and ventral pathways. *Nature Neuroscience*, 8(10), 1380–1385.
- Fasching, W. (2008). Consciousness, self-consciousness, and meditation. *Phenomenology and the Cognitive Sciences*, 7(4), 463–483.
- Forman, R., K. C. (1999). *Mysticism, Mind, Consciousness*. Suny Press.
- Frankish, K. (2017). *Illusionism: as a Theory of Consciousness*. Imprint Academic.
- Gamma, A., & Metzinger, T. (2021). The Minimal Phenomenal Experience questionnaire (MPE-92M): Towards a phenomenological profile of “pure awareness” experiences in meditators. *PLoS ONE*, 16(7), e0253694.
- Gertler, B. (2001). Introspecting phenomenal states. *Philosophy and Phenomenological Research*, 63(2), 305–328.
- Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, 15(1), 20–25.
- Gupta, B. (1998). *The Disinterested Witness: A Fragment of the Advaita Vedanta Phenomenology*. Northwestern University Press.
- Harman, G. (1990). The intrinsic quality of experience. In J. Tomberlin (Ed.), *Philosophical Perspectives 4: Action Theory and Philosophy of Mind* (pp. 31–52). Ridgewell Publishing Company.
- Ho, C.-H. (2007). Consciousness and self-awareness. *Asian Philosophy*, 17(3), 213–230.
- Hofmann, F. (2014). Non-conceptual knowledge. *Philosophical Issues*, 24(1), 184–208.
- Horgan, T., & Kriegel, U. (2007). Phenomenal epistemology: What is consciousness that we may know it so well? *Philosophical Issues*, 17(1), 123–144.
- Johnston, M. (2004). The Obscure Object of Hallucination. *Philosophical Studies*, 120(1), 113–183.
- Kammerer, F. (2018). Can you believe it? Illusionism and the illusion meta-problem. *Philosophical Psychology*, 31(1), 44–67.
- Kammerer, F. (2022). How can you be so sure? Illusionism and the obviousness of phenomenal consciousness. *Philosophical Studies*, 179(9), 2845–2867.
- Kennedy, M. (2009). Heirs of nothing: The implications of transparency. *Philosophy and Phenomenological Research*, 79(3), 574–604.
- Kind, A. (2003). What's so transparent about transparency. *Philosophical Studies*, 115(3), 225–244.
- Kriegel, U. (2009). *Subjective Consciousness: A self-Representational Theory*. Oxford University Press.
- Kriegel, U., & Williford, K. (Eds.). (2006). *Self-Representational Approaches to Consciousness*. MIT Press.
- Lycan, W. G. (2001). The case for phenomenal externalism. *Noûs*, 35(s15), 17–35.
- Mackenzie, M. D. (2007). The illumination of consciousness: Approaches to self-awareness in the Indian and Western traditions. *Philosophy East and West*, 57(1), 40–62.
- Martin, M. G. F. (2002). The Transparency of Experience. *Mind and Language*, 17(4), 376–425.
- Metzinger, T. (2020). Minimal phenomenal experience: meditation, tonic alertness, and the phenomenology of “pure” consciousness. *Philosophy and the Mind Sciences*, 1(1), 1–44.
- Montague, M. (2016). *The Given: Experience and its Content*. Oxford University Press.
- Moore, G. E. (1903). The refutation of idealism. *Mind*, 12, 433–453.
- Pace, M. (2007). Blurred vision and the transparency of experience. *Pacific Philosophical Quarterly*, 88(3), 328–354.

- Padmasambhava, Coleman., G., & Jinpa., T. (2006). *The Tibetan Book of the Dead*. (G. Dorje, Trans). Penguin Books.
- Pryor, J. (2000). The skeptic and the dogmatist. *Nous*, 34(4), 517–549.
- Ramm, B. J. (2023). Pure awareness experience. *Inquiry*, 66(3), 394–416.
- Ram-Prasad, C. (2011). Situating the elusive self of Advaita Vedanta. In M. Siderits, E. Thompson, & D. Zahavi (Eds.), *Self, No Self? Perspectives from Analytical, Phenomenological, and Indian Traditions* (pp. 217–238). Oxford University Press.
- Russell, B. (1912). *The Problems of Philosophy* (p. 1997). Oxford University Press.
- Shear, J., & Jevning, R. (1999). Pure consciousness: Scientific exploration of meditation techniques. *Journal of Consciousness Studies*, 6(2–3), 189–209.
- Siderits, M. (2011). Buddhas as Zombies: A Buddhist Reduction of Subjectivity. In M. Siderits, E. Thompson, & D. Zahavi (Eds.), *Self, No Self? Perspectives from Analytical, Phenomenological, and Indian Traditions* (pp. 308–331). Oxford University Press.
- Siderits, M., Thompson, E., & Zahavi, D. (2011). *Self, No Self? Perspectives from Analytical, Phenomenological, and Indian Traditions*. Oxford University Press.
- Siewert, C. (2004). Is experience transparent? *Philosophical Studies*, 117(1–2), 15–41.
- Smithies, D. (2012). A simple theory of introspection. In D. Smithies & D. Stoljar (Eds.), *Introspection and Consciousness* (pp. 259–293). Oxford University Press.
- Smithies, D. (2014). The phenomenal basis of epistemic justification. In M. Sprevak & J. Kallestrup (Eds.), *New Waves in Philosophy of Mind* (pp. 98–124). Palgrave Macmillan.
- Smithies, D. (2019). *The Epistemic Role of Consciousness*. Oxford University Press.
- Stoljar, D. (2004). The argument from diaphanousness. *Canadian Journal of Philosophy Supplementary*, 30, 341–390.
- Stoljar, D. (2012). Knowledge of perception. In D. Smithies & D. Stoljar (Eds.), *Introspection and Consciousness* (pp. 65–89). Oxford University Press.
- Thompson, B. (2008). Representationalism and the argument from hallucination. *Pacific Philosophical Quarterly*, 89(3), 384–412.
- Thompson, E. (2007). *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. Harvard University Press.
- Thompson, E. (2014). *Waking, Dreaming, Being: Self and Consciousness in Neuroscience, Meditation, and Philosophy*. Columbia University Press.
- Tye, M. (1995). *Ten Problems of Consciousness*. MIT Press.
- Tye, M. (2000). *Consciousness, Color, and Content*. MIT Press.
- Tye, M. (2009). *Consciousness Revisited*. MIT Press.
- Vogel, J. (1990). Cartesian skepticism and inference to the best explanation. *Journal of Philosophy*, 87(11), 658–666.
- Wright, C. (2004). Warrant for nothing (and foundations for free)? *Proceedings of the Aristotelian Society, Supplementary*, 78, 167–212.
- Zahavi, D. (2005). *Subjectivity and Selfhood: Investigating the First-Person Perspective*. MIT press.
- Zahavi, D., & Kriegel, U. (2015). For-me-ness: What it is and what it is not. In D. O. Dahlstrom, A. Elpidorou, & W. Hopp (Eds.), *Philosophy of Mind and Phenomenology: Conceptual and Empirical Approaches* (pp. 36–53). Routledge.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.