Replies to critics

Declan Smithies

Received: 26 February 2022 / Accepted: 29 March 2022 / Published online: 9 April 2022
© The Author(s), under exclusive licence to Springer Nature B.V. 2022

Abstract
I reply to my critics in this symposium on my book, The Epistemic Role of Consciousness (Oxford University Press, 2019).

Keyword Epistemic justification; Phenomenal consciousness; Representational content; Blindsight

1 Kengo Miyazono

Kengo Miyazono challenges my argument from blindsight, which is reproduced below:

(1) Blindsight subjects are not disposed to form beliefs noninferentially on the basis of unconscious visual information, but rather to withhold belief instead.
(2) Blindsight subjects are not thereby any less than fully rational.
(3) If unconscious visual information in blindsight provides a source of noninferential justification for belief, then blindsighted subjects are less than fully rational insofar as they are not disposed to form beliefs noninferentially on the basis of unconscious visual information, but rather to withhold belief instead.
(4) Therefore, unconscious visual information in blindsight does not provide a source of noninferential justification for belief. (Smithies, 2019: 81)

Miyazono’s concern is that the argument equivocates between empirical claims about blindsight and mere stipulations, and between ideal and non-ideal conceptions of rationality, in ways that undermine the force of the argument. Although I disagree, I welcome this opportunity to clarify how the argument is supposed to work.

I make the following empirical assumptions about standard (type 1) cases of blindsight. First, blindsighted patients have no conscious visual experience of...
objects in the blind field. Second, they have unconscious visual information about the blind field, which explains their ability to make reliable guesses under forced choice conditions. And third, they do not believe their guesses until they learn about their reliability, which comes to them as a surprise.

All these claims are part of the orthodox interpretation of blindsight (Weiskrantz, 1997). Of course, this orthodox interpretation is not beyond dispute. It is not my goal, however, to resolve empirical debates about the proper interpretation of blindsight. Moreover, my argument does not rely on how these debates should be resolved. Instead, we can treat the orthodox interpretation of blindsight as a thought experiment. One problem with relying on thought experiments in philosophy is that their conceivability is sometimes disputed. But when they are constructed from mainstream scientific interpretations of empirical phenomena, their conceivability is harder to deny. For my purposes, it does not matter whether the orthodox interpretation of blindsight is actually true so long as it is coherently conceivable.

My argument applies epistemic intuitions and principles to actual or counterfactual cases that accord with the orthodox empirical interpretation of blindsight. The first premise can be treated as a stipulation that blindsighted subjects do not believe their guesses until they discover their reliable track record. The second premise records the intuition that these blindsighted subjects do not thereby violate any requirement of epistemic rationality. The third premise articulates an instance of the general principle that epistemic rationality requires responding appropriately to your justifying evidence. And the conclusion follows deductively from these three premises.

In the book, I address the objection that we can explain why it is rational for blindsighted subjects to withhold belief about the blind field by invoking defeaters. On this view, unconscious vision provides defeasible justification for beliefs about the blind field, although it is defeated by background evidence, e.g., that you are unreliable. There is no guarantee, however, that such defeaters will always be available. For example, we can imagine blindsight patients who have no evidence about their own reliability or unreliability. In the absence of defeaters, they withhold belief about the blind field, but they are not thereby any less than fully rational.

Miyazono asks why we should accept these premises once we add the stipulation that our blindsight patient has no evidence about their own reliability. To begin with the first premise, why not suppose they will form beliefs about the blind field in the absence of defeaters? This is to suppose they have some default inclination to believe the contents of unconscious vision, which is inhibited by evidence of unreliability. But this is implausible on empirical grounds. There is some evidence that blindsighted patients can respond directly to the contents of unconscious vision in performing basic actions, such as pointing, grasping, and navigating around obstacles. However, there is no evidence that they are directly responsive to the contents of unconscious vision in forming beliefs about the blind field. On the contrary, they form such beliefs only by making inferences from premises about the reliability of their own guesses.

Once again, we can treat this as a mere stipulation, since I am drawing on the orthodox interpretation of blindsight to construct the thought experiment. But now Miyazono asks why we should accept the second premise that blindsight
patients are rational in withholding beliefs about the blind field. My claim here is that blindsight patients are no less rational than we are when we withhold belief about things outside our field of conscious vision. I see no plausibility in the claim that blindsight patients violate requirements of rationality merely because their beliefs are not directly responsive to the contents of unconscious vision. This is no more plausible than the idea that we violate rational requirements when we cannot make explicit the rules and representations that are implicitly represented in modular systems for visual or linguistic processing. It is highly revisionary to suppose that rationality requires our beliefs to be directly sensitive to the contents of unconscious mental representations that are in principle inaccessible to consciousness. My strategy is to use this epistemic intuition to support a theory that excludes these consciously inaccessible mental representations from our justifying evidence.

Miyazono suggests that my argument equivocates on different conceptions of rationality. At various points in the book, I invoke a distinction between ideal and non-ideal standards of rationality. In this argument, however, my talk of “full rationality” signals that I am operating with ideal standards of rationality. Withholding belief about the blind field is not just rational in some non-ideal sense that takes our cognitive limitations into account. It is no departure from ideal standards of rationality at all. An ideally rational agent with blindsight would withhold beliefs about the blind field. Of course, we are not ideally rational agents: our rationality is imperfect. But our rational imperfections reveal themselves elsewhere. There is no rational imperfection in our failure to believe the contents of unconscious vision. To my mind, this is hard to deny.

Miyazono suggests that meta-cognition may be required for rationality. I remain agnostic about whether meta-representation is involved in the subdoxastic information-processing that underpins reasoning or consciousness. Nevertheless, I explicitly reject the view that rationality requires higher-order beliefs about your own epistemic status (2019: Ch. 8). Epistemic internalism is often criticized on the grounds that it over-intellectualizes epistemic phenomena, including knowledge and rationality. One goal of the book is to develop a form of epistemic internalism that avoids this over-intellectualization problem.

A background assumption of my argument is that perceptual experience justifies belief about the external world in a way that is immediate in the weak sense that it does not depend on a posteriori justification to believe anything else. This justification can be defeated by evidence that your experience is unreliable, but it does not require evidence that your experience is reliable. After all, we cannot acquire such evidence that experience is reliable except through experience. To avoid skepticism, we must assume that experience can justify belief in the absence of evidence about its reliability. The argument from blindsight is designed to show that the same is not true of unconscious visual representation. You have no justification to believe the contents that are represented unconsciously in the visual system. That is why blindsight patients are no less than fully rational in withholding belief about the blind field.
2 Takuya Niikawa and Yasushi Ogusa

Takuya Niikawa and Yasushi Ogusa raise two challenges for the argument below:

(1) **Representationalism about perceptual experience**: Every perceptual experience has phenomenal character that is identical with the property of representing some content with presentational force.

(2) **The content principle**: Every experience that represents that p with presentational force thereby provides immediate, defeasible justification to believe that p. Therefore,

(3) **The phenomenal sufficiency thesis**: Every perceptual experience provides immediate, defeasible justification to believe some content in virtue of its phenomenal character alone (2019: 91–2).

Both challenges seek to undermine the phenomenal sufficiency thesis by targeting the content principle from which it is derived.

Their first challenge invokes a case of amnesia in which someone loses her knowledge that perceptual experience is “informative.” But exactly what does this mean? I will consider two interpretations of the case. On the first, she loses her knowledge that perceptual experience provides reliable information about the external world. On the second, she loses her knowledge that perceptual experience carries any informational or representational content at all.

On the first version of the case, the amnesic subject loses her knowledge — and, presumably, her evidence — that perceptual experience supplies reliable information about the external world. We are not supposing she has evidence that her experience is unreliable, since that would be a defeater. Instead, we are supposing she has no relevant memories that supply evidence one way or the other about the reliability of perceptual experience. She is in the same predicament as David Lewis’s *superbaby* who has no empirical information about the world.

Imagine our amnesic subject has a perceptual experience which represents some content with presentational force. Does she have any justification to believe this content? I maintain that she does. In the absence of defeaters, she has justification to believe that things are how they are represented in the contents of perceptual experience. She has this justification solely in virtue of the phenomenal character of her experience. If she acquires justification to believe that her experience is reliable, this may increase her degree of justification. But her experience alone is enough to provide some degree of justification without any independent evidence that her experience is reliable.

Perhaps Niikawa and Ogusa will insist that perceptual experience justifies beliefs about the external world only if you have justification to believe that perceptual experience is reliable. However, this requirement leads directly to skepticism. After all, where is this justification supposed to come from? According to Niikawa and Ogusa, it is not a priori but a posteriori: it has its source in perceptual experience. But how can perceptual experience justify beliefs about
reliability if we already need this justification to be in place before perceptual experience can justify believing anything at all?

On the second version of the case, the amnesic subject loses her knowledge that perceptual experience represents the external world at all. On my view, however, she cannot lose her evidence that perceptual experience represents the external world, since this evidence is supplied by perceptual experience itself. I endorse a version of representationalism on which the phenomenal character of perceptual experience is identical with a way of representing the external world with presentational force (2019: Ch. 2). And I combine this with a simple theory of introspection, according to which any experience provides conclusive introspective evidence that you have that experience (2019: Ch. 5). Hence, you cannot have a perceptual experience without thereby having conclusive introspective evidence that your perceptual experience represents the external world. The amnesic subject may forget how to use this evidence in forming doxastically justified beliefs, but her experience remains a source of propositional justification for beliefs about the external world and about her own experience.

Niikawa and Ogusa claim that perceptual experience gives you justifying evidence only if you can “appeal to the perceptual experience as evidence.” However, I explicitly deny that higher-order reflection constrains your evidence. I defend a phenomenal conception of evidence, according to which your evidence is constituted by phenomenally individuated facts about your current mental states (2019: Ch. 6). Perceptual experience gives you evidence in virtue of its phenomenal character, rather than your higher-order beliefs about your own experience. Although I endorse higher-order constraints on justification, such as the JJ principle, these are formulated in terms of propositional justification, rather than doxastic justification, to avoid standard problems of over-intellectualization and infinite regress (2019: Ch. 8). One of the main goals of the book is articulate a version of epistemic internalism that avoids these problems.

Their second challenge appeals to cases of recognitional knowledge. For example, an expert birdwatcher can know that the bird she sees is a scarlet tanager based on how it looks. In contrast, a novice birdwatcher cannot acquire this knowledge based on how the bird looks, since she does not know what scarlet tanagers look like (2019: 118). Does this example pose a problem for the content principle? Niikawa and Ogusa think so, but I disagree.

The content principle is neutral in the debate between “rich” and “thin” views about the contents of perceptual experience: it makes no claim about which properties of the external world are represented in the contents of perceptual experience. If the contents of your perceptual experience are rich enough to represent that the bird is a scarlet tanager, then you thereby have immediate justification to believe it is a scarlet tanager. Some claim that this explains the phenomenal contrast between expert and novice perception (Siegel, 2005), but I remain doubtful. An expert who misclassifies a realistic waxwork as a scarlet tanager is not thereby subject to any kind of visual illusion. The phenomenal contrast between experts and novices is more plausibly explained by patterns of attention, since experts know where to look. In principle, however, a novice can be trained to attend to the same properties as experts without yet knowing that these are characteristic of scarlet tanagers.
I have more sympathy with Matthew McGrath’s (2017) view that the expert’s recognitional knowledge depends on inference from how things look. Her inference need not be conscious. After all, she has learned to make the inference unconsciously and automatically. Nevertheless, her recognitional knowledge is inferential insofar as it depends causally and epistemically on her knowledge about how things look. If she did not know what scarlet tanagers look like, then she would not know that the bird is a scarlet tanager. Hence, she knows that the bird is a scarlet tanager only because she knows that it looks the way scarlet tanagers look. In other words, her knowledge has the following inferential structure:

(1) This looks W.
(2) If something looks W, then it is a scarlet tanager.
(3) So, this is a scarlet tanager.

Moreover, her justification to believe the conclusion has the same inferential structure. Her experience of the bird provides immediate justification for beliefs about how the bird looks, but it does not provide immediate justification to believe that it is a scarlet tanager. Given the content principle, it follows that her experience does not represent that it is a scarlet tanager. On this view, the expert and the novice may share the same experience, but there is no difference in what their experience gives them immediate justification to believe. So, this example presents no problem for the content principle.

Can McGrath’s argument be extended to show that even our knowledge of color and shape has an inferential structure? Perhaps we can know by sight that something is red and round only because we know how it looks. If so, then the content principle forces us to embrace an extremely thin view on which only appearance properties are represented in the contents of perceptual experience. McGrath (2017: 35–6) considers this view, and it is worth taking seriously. Since appearance properties are objective properties of mind-independent objects, this is not to deny that perceptual experience represents the external world. Still, I am not yet persuaded that this view is forced upon us. It may be that we know what red things look like only because we know which things are red. As far as I am concerned, this question remains wide open. In any case, the content principle is compatible with a range of different views about which properties are represented in the contents of perceptual experience.

3 Lu Teng

Lu Teng raises challenges for various epistemic principles I defend in connection with the epistemology of perception, cognition, and introspection.

Her first challenge concerns the content principle, which states that every experience that represents its content with presentational force thereby provides immediate, defeasible justification to believe its content. I argue that the scope of
this principle includes perceptual experience, but not cognitive experience, since only the former represents its content with presentational force. Teng questions whether this view is coherent and well-motivated.

What is *presentational force*? I give an ostensive definition: it is the distinctive kind of phenomenal character present in paradigm cases of sensory perception, such as seeing or hallucinating a white cube. By contrast, this phenomenal character is absent from various experiences that match in content, such as visualizing a white cube or judging that you see one. This means we cannot characterize presentational force as a special kind of *content*: for example, having an experience in which it seems that you are presented with things that make your experience true. After all, we can judge these contents to be true without thereby representing them with presentational force. Presentation is a species of force, rather than content.

Some epistemologists claim there are cognitive experiences, such as *intuitions*, that have presentational force. If presentational force is defined by ostension, however, then we must explain how to extend it beyond the paradigm case of sensory perception. What does it mean to say that intuition has presentational force? It cannot just mean that there is some phenomenal similarity between perception and intuition. On the view that I am defending, presentational force plays a foundational epistemic role in justifying beliefs without standing in need of justification. To say that intuition has presentational force is to say that its phenomenal character is similar enough to perception in the right kind of way to play this foundational epistemic role. This epistemic claim cannot be based on introspection alone. Introspection reveals various phenomenal similarities and differences between intuition and perception, but it does not settle questions about their epistemic significance. To settle these questions, we need epistemic arguments, rather than introspective ones. I argue on epistemic grounds that intuition has the wrong kind of phenomenal character to play a foundational epistemic role (2019: Ch. 12).

My own view is that only perceptual experiences and perceptual memories have presentational force. Hence, I regard presentational force as a determinate kind of sensory experience. Teng has a very different view on which presentational force is an epistemic feeling that is generated by metacognitive processes that monitor the sources of experience. I agree that we sometimes have epistemic feelings too, but that is not what I mean by presentational force. On my view, epistemic feelings cannot play the foundational epistemic role of perceptual experience.

Let me illustrate my view in connection with Teng’s example of auditory-verbal hallucinations in schizophrenia. I am agnostic about the empirical details, but let us assume that these hallucinations are generated by subdoxastic meta-representational processes that locate the cause of the experience in the external world. This hypothesis leaves open two possibilities. One possibility is that the hallucinations have exactly the same sensory phenomenal character as perceptual experience. If so, then I claim that they play a foundational epistemic role, since they represent the external world with presentational force. Another possibility, which seems rather more likely to me, is that schizophrenic patients misidentify their epistemic feelings as instances of sensory perception. In that case, their hallucinations do not have the same sensory phenomenal character as perceptual experience, although the patient...
mistakenly thinks they do. On my view, these experiences cannot play a foundational epistemic role.

Teng’s second challenge concerns doxastic conservatism, which states that if you believe that \( p \), then you thereby have defeasible justification to believe that \( p \) (2019: 117). Teng proposes a counterexample in which you are motivated by wishful thinking to believe without any evidence that it is raining in Shanghai today. In this case, she argues, you have no defeaters, since your evidence is neutral about the content of your belief, and you have no evidence about its problematic etiology. Still, your belief is unjustified. How can doxastic conservatism explain this?

We must distinguish sharply between propositional and doxastic justification. When you originally form the belief, it is doxastically unjustified because it is based on wishful thinking. Nevertheless, doxastically unjustified beliefs can give you propositional justification to retain those beliefs in the absence of defeaters. This means that beliefs that were doxastically unjustified when originally formed can become doxastically justified when retained over time so long as they are no longer held on the same basis (cf. Feldman & Conee, 2001: 8–10).

For illustration, imagine that you believe for many years that it rained in Shanghai on this particular day. Moreover, although your belief was originally motivated by wishful thinking, this is no longer a contributing factor in the retention of your belief. You are simply operating on the default assumption that your beliefs are reliable in the absence of specific reasons for doubt. Moreover, we are imagining that you have no defeaters. So, when you recall many years later that it was raining in Shanghai on the day in question, your belief is now doxastically justified, although it was doxastically unjustified when originally formed. Given its etiology, however, your belief is not knowledge even if it happens to be true. This is a Gettier case.

Teng’s third challenge concerns the simple theory of introspection, which says that any experience provides immediate and indefeasible justification to believe that you have the experience (Smithies, 2019: Ch. 4). I argue that we need this theory to explain the irrationality of epistemic akrasia. Suppose your visual experience represents that the object before you is a white cube. Given the content principle, you have justification to believe that it is a white cube in the absence of defeaters. If the simple theory is false, however, you might lack introspective justification to believe that your visual experience represents that it is a white cube. In that case, you lack higher-order justification to believe that your experience justifies believing it is a white cube. Hence, rejecting the simple theory leaves open the unattractive possibility that epistemic akrasia is sometimes justified.

As Teng notes, this argument will not persuade anyone who does not already accept the content principle. Since I have already argued for the content principle in Ch. 3 of my book, however, my goal in Ch. 4 is to argue that it should be combined with the simple theory. These are two parts of an overall package that we need to explain the irrationality of epistemic akrasia. It is true that these two parts of the package are logically independent and so it is not incoherent to endorse one without the other. Nevertheless, it is hard to explain the irrationality of epistemic akrasia without combining them into a single package.

I agree that the overall package is more strongly supported to the extent that there is independent support for its component parts. But I give independent arguments
for the simple theory, which do not rely on the content principle (2019: 165–166). First, it is supported by reflection on examples: for instance, the feeling of pain is enough by itself to justify the belief that you feel pain. You do not need to represent that you feel pain or infer that you feel pain from independently supported premises. Second, it explains why the game of giving and asking for reasons breaks down when it comes to beliefs about your own experience. If you are challenged to give your reason for believing that you feel pain, there is no dialectically satisfying answer you can give. After all, your reason for believing that you feel pain is just that you feel pain.

4 Tony Cheng

Tony Cheng takes issue with my argument against radical externalism. I define this as the view that perceptual experience justifies believing its content only if it satisfies further externalist conditions, which fail to supervene on its phenomenal character. Different forms of radical externalism impose different external conditions on the so-called “good case” in which perceptual experience justifies believing its content. Nevertheless, all forms of radical externalism agree that there is some justificatory difference between phenomenal duplicates in the good case and the bad case.

Here is my argument against radical externalism:

(1) Subjects in the bad case form the same beliefs as subjects in the good case.
(2) If perceptual experience does not provide equal justification for belief in the good case and the bad case, then subjects in the bad case are less rational than subjects in the good case insofar as they form the same beliefs as subjects in the good case.
(3) Subjects in the bad case are no less rational than subjects in the good case. Therefore,
(4) Perceptual experience provides equal justification for belief in the good case and the bad case. (2019: 99)

Cheng concedes my first premise that subjects in the bad case form many of the same beliefs as subjects in the good case. I do not claim that they form all the same beliefs, since I acknowledge that there may be externalist conditions on the contents of some beliefs, including de re beliefs. I assume only that there is some overlap in the contents of belief between phenomenal duplicates in the good case and the bad case. Cheng does not dispute this. Moreover, he also concedes the third premise that subjects in the bad case are no less rational in forming these beliefs than subjects in the good case. Instead, he disputes the second premise: he says that subjects in the bad case are less justified, but no less rational, than subjects in the good case.

What exactly does Cheng mean when he says that subjects in the bad case are rational but unjustified? He does not explicate the distinction between justification and rationality, so I am not entirely sure which distinction he has in mind. As I use these terms, they are synonymous: “to say that a belief is justified is to say that it is
reasonable or rational” (2019: 24). I am not opposed to drawing more fine-grained normative distinctions; indeed, I myself distinguish between ideal and non-ideal standards of justification or rationality. However, such distinctions must be motivated on theoretical grounds and cannot simply be read into ordinary language.

One possibility is that our disagreement is purely terminological. When Cheng says that subjects in the bad case are rational but unjustified, perhaps he means what I mean by saying that subjects in the bad case are justified while lacking knowledge. If so, there is no substantive disagreement between us. More would need to be said, however, to articulate a genuine alternative to my view.

Another possibility is that Cheng means what others mean when they say that subjects in the bad case are unjustified but blameless. I criticize this proposal on the grounds that it fails to capture the distinction between two versions of the bad case: first, the good-bad case, in which subjects succeed in conforming their beliefs to the contents of experience, although their experiences are illusory, and second, the bad-bad case, in which subjects fail to conform their beliefs to the contents of their illusory experiences owing to some cognitive delusion. Both subjects are blameless, but this fails to capture the sense in which subjects in the good-bad case are responding to their experience in an epistemically appropriate way. The challenge for radical externalism is to explain this without conceding that their beliefs are justified.

In the book, I consider recent attempts to answer this challenge by Timothy Williamson and Maria Lasonen-Aarnio (Smithies, 2019: 101–3). In different ways, they appeal to the dispositions manifested by subjects in the good-bad case. The general idea is that although subjects in the good-bad case are unjustified, they manifest dispositions that would be justified in the good case. My main objection to this strategy is that it does not generalize far enough to include Boltzmann brains whose dispositions are not anchored to the good case. But I will not elaborate on this objection here, since Cheng does not pursue this general strategy.

Instead, Cheng alludes to the work of John McDowell. I agree that McDowell would not endorse any of the options considered so far. In fact, I am not sure his view satisfies my definition of radical externalism at all. As I interpret him, McDowell does not dispute my argument, but instead proposes an alternative explanation of its conclusion.

McDowell endorses a version of epistemological disjunctivism. He concedes that subjects in the bad case are no less rational — and, if we use the terms equivalently, no less justified — than subjects in the good case. But he maintains that what makes them rational or justified in each case is different. In the good case, it is rational to believe the contents of perceptual experience because you are in a position to acquire perceptual knowledge. In the bad case, in contrast, there is some other explanation of why it is rational to believe the contents of perceptual experience. In that sense, the explanation is disjunctive.

That is the general form of the proposal, but it remains incomplete. What explains why it is rational to believe the contents of perceptual experience in the bad case? Since you cannot have perceptual knowledge in the bad case, there must be some other explanation. Presumably, however, the explanation cannot simply invoke the phenomenal character of experience. After all, we are assuming the bad case has the same phenomenal character as the good case. If we can explain the rationality of
belief in the bad case by appeal to phenomenal character alone, then presumably the same applies in the good case. Unless more can be said, the appeal to knowledge in the good case is rendered explanatorily redundant.

McDowell’s response, as I understand it, is that the explanation of rationality in the bad case is somehow parasitic on the good case. It is rational to believe the contents of perceptual experience in the bad case only because it stands in some relevant relationship to the good case. But what is that relationship? Presumably, it cannot just be the symmetric relation of sharing the same phenomenal character, or the same epistemic justification, since this would undermine the claim that the bad case is parasitic on the good case. Instead, we need some asymmetric relation that explains the rationality of the bad case in terms of the good case, rather than vice versa.

I argue elsewhere that McDowell cannot explain the rationality of belief in the bad case by appealing to the negative criterion of indiscriminability from the good case (Smithies, 2018). The bad case is indiscriminable from the good case in the sense that subjects in the bad case cannot know that they are not in the good case. And yet the same point applies equally to rocks, coma patients, and zombies, since they cannot know anything at all. The negative criterion fails to explain the positive fact that it is rational for subjects in the bad case — unlike rocks, coma patients, and zombies — to form beliefs about the external world.

In his reply, McDowell renounces this negative criterion of indiscriminability. Instead, he endorses the following positive criterion: “What is epistemically relevant about a ‘bad case’ is that it presents itself, in the subject’s consciousness, as a ‘good case’” (2018: 106). Presumably, McDowell does not mean to suggest that perceptual experience represents the higher-order content that it is a source of perceptual knowledge. It is dubious that perceptual experience represents such higher-order epistemic propositions. A more charitable interpretation is that experience in the bad case “presents itself” as a good case in the sense that it represents contents about the external world that are true only in the good case. But now I do not see how McDowell’s explanation diverges from mine. On my own view, perceptual experience justifies belief in the good case and the bad case alike by representing contents that are true only in the good case. That by itself involves no distinctive commitment to epistemic disjunctivism.

To conclude, I am not persuaded that epistemic disjunctivism has any explanatory advantage over the account I propose in the book, although the issue deserves more extended discussion. While I am not inclined to accept McDowell’s epistemic disjunctivism, or his conceptualism about the contents of perceptual experience, there are many other foundational issues on which we agree, including perhaps the most important theme in *Mind and World*: namely, that “experience is a rational constraint on thinking” (1994: 18).

5 Thomas Raleigh

Thomas Raleigh’s excellent comments focus on three main issues: zombies, propositional versus doxastic justification, and luminosity.
Let us start with zombies. On my view, zombies can have mental representations that explain their behavior, but these mental representations do not provide them with justifying reasons for belief or action, since they are inaccessible to consciousness. In a slogan, there can be representational zombies, but no rational zombies. Raleigh contends that just as we can distinguish between rational and irrational people in our world, so we can distinguish between their rational and irrational counterparts in a parallel zombie world. However, I deny that we are drawing the same distinction in each case. It is true that some zombies behave as if they are rational, while others behave as if they are irrational, but none are rational in the sense that they believe or act for justifying reasons. I do not deny that some concept of pseudo-rationality might figure crucially in the social science of zombie behavior. Nevertheless, I maintain that there is more to our concept of rationality than its role in predicting and explaining behavior from the third-person perspective. As I argue in the book, rationality also has an important first-person dimension through its connection with phenomenal consciousness.

As Raleigh notes, zombies also raise questions about the basing relation. If zombies are possible, and their beliefs are caused in the same ways as ours, then consciousness is epiphenomenal. And if the basing relation is a causal relation, then we cannot base our beliefs on conscious reasons. So how can we avoid skepticism? I deny that our beliefs are caused in the same way as zombies. If epiphenomenalism is false, as I assume, then consciousness plays an important role in causing our beliefs, which is not true for zombies. The most we can say is that their causal structure is homomorphic with ours: for every conscious state that plays some causal role in us, there is some non-conscious state that plays a similar causal role in the zombie. Nevertheless, this causal role is multiply realized: it is occupied by consciousness in us and by something else in zombies.

In any case, I do not assume that zombies are possible. This assumption leads quickly to dualism, whereas I remain steadfastly agnostic about the metaphysics of the mind–body problem. Zombies are conceivable, but it is controversial whether this means that zombies are possible. I take no stand in that debate. I assume that consciousness plays some role in causing our beliefs and actions, but it seems entirely conceivable that something non-conscious could play the same causal role in zombies. My question is whether there is any role in our mental lives that cannot conceivably be played without consciousness. And my answer is that rationality is inconceivable without consciousness.

Let us move next to the relationship between propositional and doxastic justification. Epistemologists in the reliabilist tradition explain propositional justification in terms of the capacity to form doxastically justified beliefs, whereas I regard this as fundamentally mistaken. Since we are not ideally rational agents, we cannot always convert our propositional justification into doxastic justification. This connection holds only for ideally rational agents as articulated in the following principle:

\[
\text{The modified linking principle: Necessarily, if you are fully rational, you have sufficient propositional justification to believe that } p \text{, and you adopt some doxastic attitude towards the proposition that } p \text{, then you have a doxastically justified belief that } p. \text{ (Smithies, 2019: 110)}
\]
Raleigh puts this principle to work in raising some interesting challenges for my epistemology of perception and introspection.

The first challenge appeals to attentional limitations. I claim that when your perceptual experience represents that $p$, you thereby have propositional justification to believe that $p$ in the absence of defeaters. Given the modified linking principle, this means that any fully rational agent who considers whether $p$ will form a doxastically justified belief that $p$. Plausibly, however, the contents of your perceptual experience are rich and detailed enough that you cannot attend to them all at once. As a result, you cannot simultaneously believe everything that is represented in the contents of perceptual experience. Does this natural limitation in your capacity for attention make you any less than fully rational?

I maintain that it does. The epistemic function of attention is not to supply evidence that gives you propositional justification for beliefs about the external world, but rather to exploit evidence from the contents of perceptual experience in the formation of doxastically justified beliefs (cf. Smithies, 2019: 84–85 on access consciousness). If your attention is limited, then you are thereby limited in your rational capacity to use evidence in converting propositional justification into doxastic justification. Since attentional limitations are endemic in human psychology, we tend to operate with non-ideal standards of rationality that take them into account. Nevertheless, an ideally rational agent who is perfectly responsive to their evidence would not share these human limitations. Human nature allows for some degree of imperfect rationality, but the ideal of perfect rationality remains beyond our reach.

The second challenge appeals to phenomenal duplicates in different external conditions. Suppose you are looking at two qualitatively identical apples, Fido on the left and Fifi on the right, whereas your phenomenal twin is looking at the same two apples flipped around. I claim that, since you are phenomenal duplicates, you each have propositional justification to believe exactly the same propositions. Nevertheless, you form different de re beliefs because of your different perceptual relations to Fido and Fifi. And yet neither of you is thereby any less than fully rational. So why does this not violate the Modified Linking Principle?

My answer is that de re beliefs refer to objects under modes of presentation that are reflected in the phenomenal character of experience (cf. Smithies, 2019: 109). When I form the de re belief that Fido is green, I represent Fido as the apple on the left. When my phenomenal twin forms the de re belief that Fido is green, he represents Fido as the apple on the right. So, although we both form de re beliefs about Fido, these beliefs diverge in content. Moreover, we cannot so much as entertain the same de re propositions, since Fido and Fifi are presented to us in different spatial configurations. This means we each have propositional justification to believe some de re proposition we cannot so much as entertain. In such cases, our failure to convert propositional justification into doxastic justification is no violation of rationality. The modified linking principle is explicitly designed to allow for such cases.

The third challenge concerns my discussion of Moore’s paradox (Smithies, 2019: 172–4). I claim that there are “finkish” cases in which you can have propositional justification to believe Moorean propositions, which you cannot convert into doxastic justification, since forming the belief would destroy the evidence that gives you propositional justification in the first place. Suppose you have evidence that it will
rain tomorrow but you do not respond rationally to this evidence by believing that it will rain. In that case, you have introspective evidence that you do not believe it will rain. So your total evidence justifies believing the Moorean conjunction, “It will rain tomorrow, but I don’t believe that will happen.” If you were to believe the first conjunct, however, this would give you introspective evidence that the second conjunct is false. Hence, you cannot believe the Moorean conjunction without changing the evidence that gives you propositional justification to believe it in the first place.

According to Raleigh, you do not have propositional justification to believe the Moorean conjunction because your evidence for the conjunction is defeated. More specifically, he claims, it is defeated by the consideration that the Moorean conjunction is self-falsifying in the sense that it is false whenever you believe it. I agree that the Moorean conjunction is self-falsifying, but I deny that this constitutes any kind of evidence against the Moorean conjunction in the case where you do not believe it. There is nothing to preclude cases in which you have evidence $e_1$ that justifies believing that $p$, while also having evidence $e_2$ that justifies believing that if you believe that $p$, then $p$ is false, so long as you also have evidence $e_3$ that justifies believing that you do not believe that $p$. Perhaps a demon tells you that $p$ is true, although that will change as soon as you believe it. What is strange about this evidential predicament is that no fully rational response is possible, since you are guaranteed to disrespect either $e_1$ or $e_2$ depending on whether or not you believe that $p$. But that is the whole point of finkish evidence. There is no conflict with the modified linking principle so long as we recognize that it is built into the definition of ideally rational agents that they cannot have finkish evidence of this kind.

The final issue is luminosity. I argue that propositional justification is luminous in the sense that you are always in a position to know which propositions you have justification to believe. In contrast, doxastic justification is not luminous because you are not always in a position to know whether your beliefs are properly based on your justifying evidence. This means you can have misleading higher-order evidence about doxastic justification, but not propositional justification. In an extreme case, ideally rational agents can have misleading higher-order evidence that their beliefs are based on the irrational influence of a reason-distorting drug, when in fact they are rationally based on the evidence. Interestingly, these are cases in which ideal agents can rationally believe Moorean propositions of the form, “$p$ but I don’t know that $p$, since my belief that $p$ is unjustified” (Smithies, 2019: 335–7).

Raleigh imagines a case in which someone oscillates between ideal and non-ideal rationality because they ingest a reason-distorting drug that eventually wears off. What should they believe? The question has no univocal answer, since deontic modals are context-sensitive: it depends on whether we are talking about ideal or non-ideal rationality. By ideal standards of rationality, you should always believe what your evidence supports. In the case we are imagining, you should believe that $p$, while also believing that your belief that $p$ is irrational. By non-ideal standards, in contrast, you should “Bracket” your first-order evidence when you get higher-order evidence that you cannot respond to it rationally. So, by non-ideal standards, you should reduce your confidence that $p$ and become agnostic.

On my view, the reason-distorting drug does not change any of this. What changes is whether you are ideally rational. If you are ideally rational, then ideal
standards of rationality become more salient in epistemic evaluation since you are capable of meeting them. If you are not ideally rational, then we are more likely to evaluate you by appealing to non-ideal standards of rationality that you are capable of meeting. This accounts for any tendency to say that ideal agents should do one thing, while non-ideal agents should do another. But it is important to recognize that we are equivocating between different senses of “should.”

It is an interesting question whether non-ideal requirements of rationality are sensitive to the facts about our cognitive limitations as well as our evidence about them. I am undecided, although I am inclined to say no (Smithies, 2022). If so, then we can construct cases of the kind that Raleigh has in mind by toggling our cognitive limitations in ways that remain undetectable. This would force the concession that the requirements of non-ideal rationality are not luminous. I am willing to make that concession on the assumption that the requirements of non-ideal rationality are sensitive to facts about your cognitive limitations as well as evidence about them. What is luminous, on my view, is what you are required to believe by ideal standards of rationality, since this is determined solely by your evidence.

Acknowledgements This symposium was originally scheduled to take place in Japan in July 2020, but the coronavirus pandemic forced us to reschedule the symposium online at The Brains Blog in April 2021. Although we were unable to hold the symposium in Asia, I am delighted that it will be published in the inaugural issue of Asian Journal of Philosophy. I am grateful to all the commentators for their careful engagement with my book and especially to Takuya Niikawa, Kengo Miyazono, Tony Cheng, and Nikolaj Pedersen, for organizational help along the way.

Data availability Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

References

Feldman, R., & Conee, E. (2001). Internalism Defended. American Philosophical Quarterly, 38, 1–18.
McDowell, J. (1994). Mind and world. Harvard University Press.
McDowell, J. (2018). Response to Smithies. Analytic Philosophy, 59, 106–107.
McGrath, M. (2017). Knowing what things look like. Philosophical Review, 126, 1–41.
Siegel, S. 2005. Which properties are represented in perception? In Perceptual experience, edited by T. Gendler and J. Hawthorne. Oxford University Press, 581–503.
Smithies, D. (2018). Discussion of John McDowell’s “Perceptual experience and empirical rationality.” Analytic Philosophy, 59, 105–106.
Smithies, D. (2019). The epistemic role of consciousness. Oxford University Press.
Smithies, D. 2022. The epistemic function of higher-order evidence. In Propositional and doxastic justification: New essays on their nature and significance, edited by P. Silva and L. Oliveira. Routledge, 97–120.
Weiskrantz, L. (1997). Consciousness lost and found. Oxford University Press.