

Amber Ross

Illusionism and the Epistemological Problems Facing Phenomenal Realism

Abstract: *Illusionism about phenomenal properties has the potential to leave us with all the benefit of taking consciousness seriously and far fewer problems than those accompanying phenomenal realism. The particular problem I explore here is an epistemological puzzle that leaves the phenomenal realist with a dilemma but causes no trouble for the illusionist: how can we account for false beliefs about our own phenomenal properties? If realism is true, facts about our phenomenal properties must hold independent of our beliefs about those properties, so mistaken phenomenal beliefs must always remain an open possibility. But there is no way to identify the phenomenal facts that make these beliefs false other than by mere stipulation. If illusionism is true, then the state of affairs regarding what a subject's experience seems like is just the illusion itself; there are no further facts of the matter about which the subject might have mistaken beliefs, so the problem does not arise.*

Phenomenal realism, whether radical or conservative, appears to be the default position regarding conscious experience. Anything less than full-fledged realism seems to deny the existence of our subjective inner lives. But the proposal that Keith Frankish gives us here, a type of anti-realism he calls 'illusionism' about phenomenal properties, has the potential to leave us with all of the benefit of taking conscious experience seriously and much less of the unnecessary metaphysical and conceptual baggage that accompanies a realist interpretation of these properties.

Correspondence:

Amber Ross, University of Toronto, Canada. *Email:* amber.ross@utoronto.ca

In categorizing phenomenal properties as illusions, it is important to note that Frankish is not suggesting that our inner lives are any less rich than they seem to be. On the contrary, our conscious experience is exactly as rich as it seems to be, not because there are real phenomenal properties present in our experience but because being rich and full of feeling is precisely how our experience seems to us, how our introspective representational mechanisms present our experiences to us. Frankish takes issue with the characterization of a philosophical zombie — a creature physically identical to an ordinary human being — as something ‘with no inner life, whose experience is completely blindsighted’ (Frankish, this issue, p. 22). The ‘inner life’ of any creature whose introspective representational mechanisms present their experience as rich and full of feeling *just is* rich and full of feeling. According to Frankish, ‘having the kind of inner life we have... consists of having a form of introspective self-awareness that creates the illusion of a rich phenomenology’ (*ibid.*). If we take illusionism seriously, then there is no further question of whether we are accurately representing *real* (rather than ersatz or pseudo) phenomenal properties. Zombies are just as correct in their judgments about how their experience *seems to them* as their ordinary human counterparts.

As difficult as it might be to come to terms with the idea that phenomenal consciousness is a type of illusion, the epistemological problems facing the phenomenal realist may actually outweigh those facing illusionism. As Frankish aptly notes, ‘we have no introspective way of checking the accuracy of our introspective representations, and so cannot rule out the possibility that they are non-veridical’ (p. 28). He mentions this in the context of an argument for taking illusionism seriously, but there are multiple ways in which our introspective representations could turn out to be non-veridical, and potentially the most complicated arise for realist views. If illusionism is correct, then our introspective representations are non-veridical in so far as we interpret them as representing real properties rather than mere ‘intentional objects, or a sort of mental fiction’ (*ibid.*). If realism is correct, certain failures of representation must still be a possibility. However, the manner in which our introspective representations can turn out to be non-veridical is more complicated: while the existence of phenomenal properties would depend upon the existence of our conscious experience, if phenomenal properties are real then their nature is necessarily independent of our beliefs about our experience.

Those who admit phenomenal properties into their ontology encounter serious epistemic difficulties, regardless of whether their underlying metaphysics is physicalist or dualist. To use Frankish's terms, both 'radical' and 'conservative' phenomenal realists face challenging puzzles regarding our epistemic relation to the phenomenal properties of our experience. To be a realist about phenomenal properties one must be able to defend the possibility of a scenario in which a subject believes herself to have a conscious experience with a particular phenomenal character while she is actually having a conscious experience with a different phenomenal character. That is, subjects must be able to genuinely hold mistaken beliefs about the content of their own conscious experience, about *what it's like* to be them at that moment. Cases in which subjects are not closely attending to their experience, or in which they misremember previous experiences and make inaccurate comparisons between current conscious experiences and prior ones, are easy to conceive. But if phenomenal realism is true, then subjects must be able to make mistakes about the content of their conscious experiences even when they are attending carefully, and even when their beliefs only concern their current conscious experiences. If a property is real, there are objective facts about that property, which is to say that whatever is true about these properties will be true regardless of a subject's beliefs about them. So a subject must be able to hold mistaken beliefs about the intrinsic (rather than relational) content of their occurrent conscious experience. Any treatment of phenomenal properties that falls short of this will fail to qualify as genuine realism, and will rather be some sort of disguised illusionism.

Constructing a scenario in which a subject believes herself to have a conscious experience with phenomenal character Φ (that is, a conscious experience that instantiates phenomenal property Φ) while she is actually having an experience with phenomenal character Ψ is fairly difficult. Chalmers (2003) attempts to do so when discussing the epistemology of phenomenal belief, and examining the scenario he constructs will illuminate certain epistemological challenges facing the phenomenal realist, challenges an illusionist will be able to avoid. To provide a framework for conceptualizing phenomenal belief — beliefs about our own occurrent conscious experiences — Chalmers coins a set of technical terms. One is the notion of a *direct phenomenal concept*, a concept the subject deploys via introspection which 'by its nature picks out instances of an underlying demonstrated phenomenal quality' (*ibid.*, p. 242) or phenomenal property. When

Mary sees colour for the first time and has the thought, ‘Oh, so that’s *what it’s like to see red*’, the concept we could articulate as ‘what it’s like to see red’ which she deploys here is what Chalmers calls a ‘direct phenomenal concept’.¹ Direct phenomenal concepts partially constitute *direct phenomenal beliefs*, the kind of belief that ‘identifies the referent of that concept with the very demonstrated quality’ (*ibid.*), as Mary’s belief above identifies the demonstrated phenomenal quality (the referent of ‘that’) with the referent of the direct phenomenal concepts (‘what it’s like to see colour’). So a direct phenomenal belief is formed when a subject internally gestures towards an experience she is having at time t_1 with an introspective token-demonstrative ‘*this (Q)...* experience’, and this kind of belief lasts only as long as the subject attends to the experience.² If phenomenal properties are real, whether they are ultimately physical or non-physical, it should be possible for subjects to hold direct phenomenal beliefs, since these beliefs, when justified and true, constitute our phenomenal knowledge.³ And it is this sort of phenomenal belief that subjects must be able to hold in error, since the truthmaker for such a belief is not how a subject’s conscious experience seems to her but rather the facts about the phenomenal properties instantiated in her conscious

-
- 1 It would be more appropriate to characterize the direct phenomenal concept in Mary’s belief as ‘what it’s like to *have this experience*’ rather than ‘what it’s like to *see red*’, since characterizing it as a *red experience* introduces a relational element we are trying to avoid here, in particular a connection between the experience she is having and what she already knows about colour phenomena, in particular, that some stimuli evoke red responses.
 - 2 Since both direct phenomenal beliefs and direct phenomenal concepts last only as long as a subject attends to a particular experience, there are some views of concepts according to which direct phenomenal concepts will not qualify as concepts, and will leave the status of Mary’s belief an open question. For example, per the definition of ‘concept’ from Prinz (2007) and Millikan (2000), for something to be a concept it must be re-deployable; that is, it must be the kind of thing a subject can deploy on multiple occasions, as the content is determined or fixed only by similarities between various instances in which a subject is disposed to deploy the concept. So by merely attending to an experience a subject is having at time t_1 and internally gesturing toward it with an introspective token-demonstrative ‘*this (Q)...* experience’, the subject will not have formed a phenomenal concept. On the Millikan/Prinz view, Q here cannot be a concept unless it can be reused to identify multiple instances of the same phenomenal state, and this would conflict with the stipulated definition of direct phenomenal concepts.
 - 3 While there is room in logical space for a position that embraces realism about phenomenal properties but rejects the possibility of phenomenal knowledge, such a position would lack any intuitive or philosophical appeal.

experience, which are at least conceptually if not metaphysically independent of her beliefs about the character of her experience.

The ‘clear case’ of failed direct phenomenal belief that Chalmers (2003) attempts to provide is perhaps as close as one can come to presenting such a case, but it is far from obvious that it actually constitutes a genuine case of this type. A direct phenomenal belief is only formed when a subject deploys a demonstrative phenomenal concept and a direct phenomenal concept ‘based in the same act of attention’ (*ibid.*, p. 236) and lasts only as long as that act of attention. For a subject to successfully form a direct phenomenal belief, her demonstrative phenomenal concept and direct phenomenal concept must be ‘appropriately aligned’ (*ibid.*). A subject will attempt but fail to form a direct phenomenal belief (or will form a false direct phenomenal belief) in cases where her demonstrative act fails to pick out a referent. In such a case, a subject intends to attend to some phenomenal property instantiated in her phenomenal experience and (in so doing) to form a direct phenomenal concept of that property, but somehow fails in this attempt. Chalmers attributes this failure to a ‘mismatch’ between the cognitive element of the demonstration and targeted experiential element, i.e. the phenomenal property.

Chalmers’ example is the case of mildly-misfortunate Nancy, who attends to a coloured patch of her phenomenal field, acting cognitively as if to demonstrate a highly specific phenomenal shade. Nancy intends to attend to her phenomenal experience in such a way that she would demonstrate a patch of phenomenal colour; that is, the object of Nancy’s attention and intended demonstration — the patch of phenomenal colour — is a phenomenal property. But her demonstration fails, purportedly because the cognitive elements and targeted experiential elements of her attempted demonstration are ‘mismatched’. In Chalmers’ thought experiment,

...Nancy has not attended sufficiently closely to notice that the patch has a nonuniform phenomenal color: let us say it is a veridical experience of a square colored with different shades of red on its left and right side... (*ibid.*, p. 237)

If we consider this scenario for a moment, we will see that there are at least two stipulated features of Nancy’s situation that make her ‘false belief’ seem plausible, both of which are suspect.⁴ First, the notion

⁴ We might choose to say that Nancy has failed to form a direct phenomenal belief, rather than attributing to her a full-fledged false belief. For our purposes here either

that Nancy's phenomenal field could be differently shaded on its left and right side while she believes it to be uniformly coloured seems *prima facie* plausible when it is described as a veridical experience of something in the external world, the properties of which can be objectively confirmed. But we should ask what *reason* we have for positing that Nancy's phenomenally conscious experience is a veridical representation of this multi-shaded square when this assertion conflicts with the subject's own belief about how her phenomenal field seems to her. To Nancy, it seems as if her phenomenal field has a consistent colour throughout, and without any way to confirm the 'phenomenal facts' regarding her conscious experience, to claim that the phenomenal content of her experience is anything other than how she takes it to be would be mere stipulation without support. Perhaps connecting the facts about Nancy's phenomenal field to states of affairs in the external world could provide suitable support, but since we are concerned solely with her subjective phenomenal experience, any connection between her phenomenal field and the world it represents is irrelevant, as Chalmers subsequently seems to agree. A later version of this material appears in Chalmers (2010, chapters 8 and 9), which are a near reprint of Chalmers (2003). In his subsequent account of Nancy's failed phenomenal belief, the qualifier 'veridical' is omitted from the description of the content of her phenomenal experience. Instead, the passage reads: 'Let us say that she has an experience of a square colored with different shades of red on its left and right side' (2010, p. 270). And rightly so, as the use of 'veridical' is unlicensed here, an unsupportable stipulation. But if the realist cannot appeal to the notion of 'veridical representation' to specify the content of Nancy's experience, he loses the intuitively plausible grounds for his claim that the *actual* content of the subject's phenomenal experience and her *beliefs* about her phenomenal experience are 'mismatched'. Her failure is left unexplained.

Stipulating that Nancy's beliefs about her experience do not accurately represent how her experience seems to her flies in the face of our intuitive conception of our relation to our own conscious experience, a problem the illusionist can easily avoid. If illusionism is true, a subject may still misjudge a current quasi-phenomenal red

interpretation would be adequate: Nancy has made an odd mistake in judgment regarding the phenomenal properties present in her conscious experience, which has resulted in either a false belief or a failed attempt at forming such a belief.

experience as being identical to a past quasi-phenomenal red experience. But if Frankish is correct and there are no real phenomenal properties, then a subject's judgments about how her current experiences seem to her now (rather than how they relate to other experiences) cannot fail to match some fact of the matter about the *real* phenomenal properties of her experience. The realist, however, must be willing to accept that a subject's beliefs may always misrepresent the facts of the matter about their phenomenal experiences. He is committed to the notion that the facts about the phenomenal character of a subject's conscious experience are set by facts about the real phenomenal properties instantiated in her experience, and part of what makes the view *realism* is that those facts are in no way determined by a subject's beliefs about how her experience seems to her. Any view according to which the subject's beliefs about the character of her conscious experience do play a role in determining the facts of the matter about her conscious experience is a non-realist, illusionist type of view.

The second feature of this situation that makes Nancy's false belief seem plausible is that Nancy is described as not attending closely to her experience. The only circumstance Chalmers acknowledges in which a subject can attempt but fail to form a direct phenomenal belief is when the subject is not adequately attending to her experience (see Chalmers, 2003). As Frankish points out (p. 30), a radical or anti-physicalist realist may appeal to 'acquaintance' as our form of epistemic access to the phenomenal properties of our experience, so long as we are adequately attending to those properties (see Chalmers, 1996, pp. 196–7; 2003, pp. 246–54). And, as Frankish notes, the acquaintance relation is no magic bullet, even for anti-physicalists, since we must mentally represent our phenomenal properties in order to think or talk about them, and the representational process is by its nature potentially fallible (p. 30). If phenomenal realism is true, then Nancy must be able to form the false belief that her phenomenal field is consistently coloured when it is actually non-uniform, regardless of how closely she is attending to her experience. Her belief represents the character of her conscious experience, it does not constitute the character of that experience, and representation always carries the possibility of misrepresentation. Even if we grant that there is some sort of certainty guaranteed by the acquaintance relation, this certainty will not transfer to our beliefs about the phenomenal character of our conscious experience. If phenomenal properties are real, we must be capable of making errors in our judgment about their nature regardless

of how closely we attend to the content of our experience. As long as the phenomenal and cognitive components of our mental states remain distinct (either conceptually or metaphysically), the cognitive effort we put forth in forming those beliefs will provide no guarantee that our beliefs will accurately represent the facts about our phenomenal properties. But to say that Nancy attends as closely as possible to the phenomenal character of her conscious experience and still forms false beliefs about it seems to leave us at an awkward distance from the content of our own experience, a distance that the phenomenal realist is committed to embracing.

The phenomenal realist is burdened with the problem of finding a plausible way to answer the question of whether a subject's phenomenal beliefs satisfy the criteria for phenomenal knowledge. Not all phenomenal beliefs will succeed here; there must be room for some phenomenal beliefs to turn out false. Hence the realist encounters a serious *problem of one's own mind*. An illusionist has no such problem, and faces no analogous challenge; subjects hold phenomenal beliefs, but there is no further question of whether these beliefs fit the criteria for phenomenal knowledge, because according to the illusionist phenomenal beliefs do not represent real properties — there is nothing against which their 'accuracy' can appropriately be measured. Phenomenal belief is the whole story, and there is nothing to be gained by searching for further 'phenomenal facts' nor by asking whether those beliefs amount to 'phenomenal knowledge'. If the illusionist is correct, the answer to this question is 'no'; there is no such thing as phenomenal knowledge in the sense advocated by the phenomenal realist.

For the phenomenal realist, the standards against which phenomenal beliefs can be deemed true or false are set by phenomenal facts — facts about one's own phenomenal states that are independent of one's beliefs about those states. For the radical (anti-physicalist) realist, these facts will be independent of any physical facts about the subject. For the conservative realist, phenomenal facts will be a special subset of physical facts, but must still be (either conceptually or in actuality) independent of a subject's phenomenal beliefs. As Frankish rightly points out, conservative realism balances on a knife edge, and if it turns out that, though we may have phenomenal beliefs (beliefs about how our experience seems to us), there is no such thing as an actual phenomenal fact, the conservative position will collapse into illusionism.

References

- Chalmers, D.J. (1996) *The Conscious Mind: In Search of a Fundamental Theory*, New York: Oxford University Press.
- Chalmers, D.J. (2003) The content and epistemology of phenomenal belief, in Smith, Q. & Jovic, A. (eds.) *Consciousness: New Philosophical Perspectives*, pp. 220–272, Oxford: Oxford University Press.
- Chalmers, D.J. (2010) *The Character of Consciousness*, New York: Oxford University Press.
- Millikan, R. (2000) *On Clear and Confused Ideas*, Cambridge: Cambridge University Press.
- Prinz, J. (2007) Mental pointing: Phenomenal knowledge without concepts, *Journal of Consciousness Studies*, **14** (9–10), pp. 184–211.