Commentary On Chadd Kidd's "The Idols of Inner-Awareness: Towards Disjunctive Self-Representationalism"

Stephen Pearce
University of Western Ontario
spearce7@uwo.ca

Introduction

According to *extrinsic* views of phenomenal consciousness, whether a given mental state is phenomenally conscious is a matter of how it is related to *other* mental states. Higher-order representational theories (HO) hold that a lower-order mental state M is phenomenally conscious when it is represented (in the right sort of way) by a *numerically distinct* higher-order mental state M*. Self-representational theories (SR) add that the lower- and higher-order mental states must be somehow constitutively *related*.

HO and SR theories of phenomenal consciousness, then, hold that lower-order mental states have phenomenal character when they are represented, in the right sort of way, by a higher-order mental state.

In what follows, I will use "qualitative property" to refer to those properties that are had by lower-order mental states that may or may not be phenomenally conscious. As I use "phenomenal character" (and "phenomenal property"), only lower-order mental states that are represented in the right sort of way by higher-order states have phenomenal character. For example, a mental state with *qualitative* blueness may or may not be phenomenally conscious. When it is represented in the right sort of way by a higher-order state, it has *phenomenal* blueness. This is the sense in which phenomenal character is *extrinsic*: it is identical to the qualitative properties of lower-order mental states which are *appropriately related* to other mental states.

According to HO theories, then, a lower-order state M has phenomenal blueness when:

1) M has qualitative blueness.
2) A higher-order state M* represents that one is in state M with qualitative blueness.

And according to SR theories, a lower-order state M has phenomenal blueness when:

1) M has qualitative blueness.
2) A higher-order state M* represents that one is in state M with qualitative blueness.
3) M and M* are constitutively related to one another (perhaps by being parts of the same mental state)

On Chad Kidd's preferred version of SR, the constitutive relation between the lower- and higher-order state must exhibit *one-sided-dependence*:

> "The higher-order state (M*) constitutive of a phenomenally conscious experience as of S's being in M is "one-sidedly dependent" on the first-order state (M) it represents. That is:
>
> Necessarily, if a token of the state-type M* exists in w, then a token of the state-type M, which

> is the intentional object of this token of M*, exists; but, possibly, a token of M exists in w without a token of M*." (pg. 4)

In what follows, I will only consider SR theories that exhibit such one-sided-dependence between the lower- and higher-order states.

Conservative Indefeasibility

In his paper, Kidd attempts to argue for the superiority of SR over HO theories by demonstrating that only SR theories are able to accommodate the *epistemic specialness* of introspective beliefs about the phenomenal character of one's occurrent experiences. I will call such beliefs *phenomenal introspective beliefs.*

Kidd holds that phenomenal introspective beliefs are *conservatively indefeasible*:

> ".. a conservatively indefeasible belief is one whose justification is not possible to undermine without also modifying the evidence or warrant that epistemically grounds the belief." (pg. 11)

In other words, a conservatively indefeasible belief is one whose evidential ground *guarantees* its truth. Thus, one cannot undermine the truth of a conservatively indefeasible belief without altering the facts that are its evidential ground.

It is plausible that phenomenal introspective beliefs are conservatively indefeasible in this way. This is because it seems like the evidential ground of such beliefs *just is* the phenomenal character of ones occurrent experiences. When one judges, in the usual way, that one is having a phenomenally red experience, what justifies this belief is just the fact that one *is* having a phenomenally red experience.

In fact, this seems like the *only* way phenomenal introspective beliefs could be conservatively indefeasible. Since conservative indefeasibility requires that a belief's evidential ground guarantees its truth, a conservatively indefeasible belief must be one whose evidential ground is the same as its truth maker. And since the truth maker of a phenomenal introspective beliefs is the phenomenal character of one's occurrent experiences, this must be the evidential ground of phenomenal introspective beliefs if they are to be conservatively indefeasible.

In contrast to phenomenal introspective beliefs, ordinary perceptual beliefs – that is, beliefs about material objects whose evidential ground is ones current perceptual experiences – are plausibly *not* conservatively indefeasible. One's perceptual belief that there is a red cup on the table is justified by it *visually seeming* to be the case that there is a red cup on the table. But it can visually seem that there is a red cup on the table even though there is not, perhaps because of the presence of red light in one's environment. The red light, then, defeats one's perceptual belief without modifying its evidential ground.

Let's call an epistemic situation that tends to result in a true belief a 'good' case with respect to that belief, and one which tends to result in a false belief a 'bad' case. Then we can say that beliefs that are *not* conservatively indefeasible are such that the 'bad' cases and the 'good' cases share a *common evidential ground*. It might visually seem to be the case that there is a red cup on the table, whether or not there actually is. In contrast, the good and bad cases for conservatively indefeasible beliefs have distinct evidential grounds.

Kidd holds that HO theories are not able to accommodate the conservative indefeasibility of phenomenal introspective beliefs. Here is Kidd's argument (pg 12):

> (1) According to HO, it is possible for a subject S to be in a state of inner-awareness M*, which represents S as being in a state M, without S's actually being in M.
>
> (2) Also, M* and M are distinct existences, i.e., the same most specific kind of state of inner-awareness M* can exist in both standard and non-standard cases of experience.
>
> (3) Thus, since M* is a common factor of both standard and non-standard cases of inner-awareness, it is also possible for S's introspective belief that S (herself) is in M to be defeated while S is still experientially aware of herself as being in M.
>
> (4) Therefore, by definition, introspective belief about current conscious experience is not conservatively indefeasible.

Suppose that the higher-order state M* represents that one is in a lower-order state M with qualitative redness – that it, M* represents that one is having a red experience. Because HO takes the higher-order state to be distinct from the lower-order state, this is possible even if one is *not* in fact having a red experience. Thus there is a 'common-factor' between cases where one is having a green experience and cases where one is not.

This argument has a missing premise, however. In order to derive a conclusion about the nature of *introspective beliefs* about the phenomenal character of occurrent experiences, we must hold that the evidential grounds of such beliefs *are* higher-order states. It is these states of inner-awareness – that represent that one is having a certain kind of experience – that justify such introspective beliefs.

> (HOJ) When one forms a phenomenal introspective belief, in the usual way, to the effect that one is having an experience with phenomenal property P, what justifies this belief is a higher-order state, which represents one as being in a lower-state with qualitative property P.

In order for Kidd's argument to go through, then, it is not enough to say that the higher-order state is a common-factor in good and bad cases of phenomenal introspective belief. Rather, we must say that the higher-order state is a common *evidential* ground of both good and bad cases of phenomenal introspective beliefs. And this requires HOJ.

Unfortunately, HOJ is in tension with the conjunction of two claims:

> (1) The evidential ground of phenomenal introspective beliefs is the *phenomenal character* of one's occurrent experiences.
>
> (2) According to both HO theories, it is *lower-order* states – not higher-order states – that have phenomenal character.

This second claim is the essence of extrinsic views on phenomenal consciousness: that a mental state is phenomenally conscious when it is related in the right sorts of ways to *other* mental states. HO theories hold that lower-order states are phenomenally conscious when they are represented, in the right sort of way, by a distinct higher-order state. Thus if HOJ is true, then phenomenal character cannot be the

evidential ground of phenomenal introspective beliefs, and thus they cannot be conservatively indefeasible.

In order to demonstrate that HO theories cannot accommodate the conservative indefeasibility of phenomenal introspective beliefs, we might hold that it is *higher-order states* that have phenomenal character. On this view, a state has phenomenal character when it has the right sort of *higher-order content*: namely, when it represents that one is in *another* state with a certain qualitative character. But this is to abandon an extrinsic view of phenomenal consciousness.

For Kidd's argument to go through, then, it looks as though we need to deny (1). We must reject that it the phenomenal character of one's occurrent experiences that is the evidential ground of introspective beliefs about them. Instead, we must hold that it is those mental states in virtue of which we are *aware* of our phenomenal experiences that justifies such phenomenal introspective beliefs. In other words, we must hold that the evidential ground of phenomenal introspective beliefs are the *higher-order* mental states that represent that we are having a certain kind of experience.

On this view, my introspective belief that I am having a phenomenally red experience is not justified by the phenomenal red of my current experience. Rather, it is justified by my higher-order state that represents that I am currently having a red experience. On this view, HO theories *do* posit a common *evidential* ground for good and bad cases of phenomenal introspective belief. I may be in a higher-order state that represents that I am currently having a red experience, even though I am not. Thus, phenomenal introspective belief is not conservatively indefeasible.

Why think that phenomenal introspective beliefs *could* be conservatively indefeasible once we reject that their evidential ground is the phenomenal character of one's occurrent experiences? As we saw above, it looks as though the only sorts of beliefs that could be conservatively indefeasible are those whose truth makers are identical to their evidential ground. But we have seen that we must take these to be distinct in the case of phenomenal introspective beliefs. Their truth makers are phenomenally conscious mental states, and those are lower-order states which are appropriately represented by higher-order states. But their evidential grounds are the higher-order states.

But we were too quick to claim that conservatively indefeasible beliefs require their truth makers to be *identical* to their evidential ground. Instead, conservatively indefeasible beliefs can have truth makers that are *metaphysically necessitated* by their evidential ground. This is strong enough to get what we need: that you cannot defeat such beliefs without changing their evidential ground.

And this is precisely the sort of relationship that SR theories posit between the relevant higher-order and lower-order states. According to SR theories:

> It is impossible for S to be in a state of inner-awareness M*, which represents S as being in a state M, without S's actually being in M. (pg. 12)

In other words, according to SR theories, it is impossible to be aware of myself as having a phenomenally blue experience without actually having a phenomenally blue experience. Thus, even though my introspective belief that I am currently having a blue experience is *not* my having a blue experience, what it *is* grounded on – my being aware of myself as having a blue experience – metaphysically necessitates that I *am* having a blue experience. Thus, according to SR theories, phenomenal introspective beliefs are conservatively indefeasible.

Fallible Inner-Awareness and SR Disjunctivism

It looks, then, as though SR theories – but not HO theories – are able to accommodate the conservative indefeasibility of phenomenal introspective belief. But as Kidd notes, this seems to come at a cost. It looks like SR theories are incompatible with the following thesis:

> *Infallibility of inner-awareness (FIA)*: It is possible for it to seem to you as if you are having a certain kind of experience without actually having an experience of that kind.

If we understand "seeming to have a certain kind of experience" as being in a state of inner-awareness that represents that one is having a certain kind of experience, then this is clearly incompatible with SR. FIA entails that one can be in a state of inner-awareness M*, which represents one as being in a state M, without one actually being in M.

To see the plausibility of FIA, Kidd discusses the phenomenon of false pains due to "dental fear". In some cases, subjects report that they feel pain even after anesthetics have been administered. If we are to take subjects' reports at face value, it looks as though this is a case where it seems to them that they are in pain – that is, having an experience with a certain kind of phenomenal character – even though they are not. The possibility of false pains, then, compels us to accept FIA.

FIA entails that good and bad cases of inner-awareness are introspectively indistinguishable. In the normal case, according to SR, your inner-awareness of yourself as being in pain necessitates that you are in fact in pain. But in bad cases, such as false pains due to dental fear, you can be in a state of inner-awareness introspectively indistinguishable from that of a good case, even though you are not in fact in pain. This suggests an argument from FIA to the falsity of SR:

> 1) In most cases, if you are in a state of inner-awareness M*, which represents you as being in a state M (in the right sort of way), then you are actually in M.
>
> 2) It is possible to be in a state of inner awareness B* that is introspectively indistinguishable from M*, even though you are not in M.
>
> 3) If B* and M* are introspectively indistinguishable, then they have the same nature.
>
> 4) Therefore B* represents you as being in a state M (in the right sort of way.)
>
> 5) Therefore it is possible to be in a state of inner-awareness M*, which represents you as being in a state M (in the right sort of way), without actually being in M.
>
> 6) Therefore SR is false.

Thus the possibility of false pains, coupled with the principle (in premise 3) that pairs of introspectively indistinguishable mental states have the same nature, results in the negation of SR.

The way out for the SR theorist, according to Kidd, is to reject premise 3. On this view, while good cases of inner-awareness awareness guarantee that one is having the appropriate experience, bad cases do not, *even though these cases are introspectively indistinguishable*.

This is to adopt a *disjunctivist* view of inner-awareness. A state of inner-awareness to the effect that it

seems to me that I am having a pain experience, for example, is *either* a state that *factively* represents (in the right sort of way) that I am having a pain experience, *or* a state that is *merely introspectively indistinguishable* from such a factive representation. Call the former the 'standard' cases, and the latter the 'non-standard' cases.

Here is how Kidd characterizes disjunctivism about inner-awareness:

> *Disjunctivism (DISJ)* An experience that is, from S's point of view, as of S (herself) being in M is either a case of S's actually being in M or merely seeming to be in M. (pg. 13)

It looks as though the SR theorist who adopts such a disjunctivist view of inner-awareness can have the best of both worlds. She can accommodate the conservative indefeasibly of phenomenal introspective belief. When such introspective beliefs are grounded on standard cases of inner-awareness, their truth *is guaranteed.* And she can allow for the fallibility of inner-awareness, such as the possibility of false pain. This is because non-standard cases of inner-awareness do *not* guarantee the truth of phenomenal introspective beliefs about them – they are merely introspectively indistinguishable from those that do.

Unfortunately, things are not so simple. Adopting a disjunctivist view about inner-awareness requires that we reject that states of inner-awareness are the evidential ground of phenomenal introspective beliefs.

To get there, we need the following thesis:

> *Introspective Evidence Indistinguishability (IEI):* If two states are introspectively indistinguishable, then they play the same role in justifying introspective beliefs.

According to IEI, if two states A and B are introspectively indistinguishable, then A justifies an introspective belief *if and only if* B does. Compare IEI to a similar thesis about perception:

> *Perceptual Evidence Indistinguishability (PEI)*: If two objects are perceptual indistinguishable, then they play the same role in justifying perceptual beliefs.

Consider the following two cases. In case A, you see a dog down the road. In case B you see what is in fact a cardboard cutout of a dog down the road. The cutout in case B is so lifelike that it is perceptually indistinguishable (from where you are standing) from the real dog.

In case A, your perception of the dog justifies you in believing that there is a dog down the road. But it is plausible that in case B your perception of the cardboard cutout *also* justifies you, absent relevant defeaters, in believing that there is a dog down the road. Why? Because these two cases are perceptually indistinguishable.

What should we say is the *evidential ground* of the perceptual belief that there is a dog down the road? What we *cannot* say is that it is a *veridical* perception of a dog down the road. After all, your perception as of a dog down the road does *not* need to be veridical in order to justify your belief that there is a dog down the road. This belief, then, is *not* evidentially grounded on your perception *being veridical*.

We might say that the evidential ground of this perceptual belief is: either perceiving a dog down the road, or perceiving something perceptually indistinguishable from a dog down the road. But this is just

to say that the perceptual belief that there is a dog down the road is evidentially grounded by it *perceptually seeming* to be the case that there is a dog down the road.

This is why perceptual beliefs are not conservatively indefeasible: there is a common evidential ground between good and bad cases. What justifies your perceptual belief that there is a dog down the road is it perceptually seeming to be the case that there is a dog down the road, and this is possible whether or not there is *in fact* a dog down the road.

The analogy to introspection should be clear. Consider the following two cases. In case A, I am having a pain experience, and am in a state of inner-awareness that factively represents that that I am having a pain experience. In case B, I am *not* having a pain experience, but am in a state of inner-awareness introspectively indistinguishable from that in case A.

In case A, your state of inner-awareness clearly justifies my belief that I am having a pain experience. This is a 'standard' case of phenomenal introspective belief, according to the SR theorist. But, I want to say, it is plausible that in case B, my state of inner-awareness *also* justifies – to the same degree – my belief that I am having a pain experience. Why? Because the two cases are introspectively indistinguishable.

If non-standard cases of introspective phenomenal belief are possible, then we cannot say that the evidential ground of such beliefs is a *factive* state of inner-awareness. My phenomenal introspective belief that I am having a pain experience is equally justified in both standard and non-standard cases from the perspective of introspection. So the evidential ground of such beliefs must be that which is *common* to both of these cases.

According to the disjunctivist about inner-awareness, standard and non-standard cases of inner-awareness correspond to different *kinds* of states. In the standard case, the state of inner-awareness is a higher-order state which necessarily veridically represents that one is having a certain kind of experience. In the non-standard cases, one is merely in a state that is introspectively indistinguishable from such a factive higher-order state.

But as we can see from DISJ, the states of inner-awareness operative in standard and non-standard cases *do* fall under a common kind: that is, states according to which it *seems* to one as if one is having a certain kind of experience. This is the sense in which these two states are introspectively indistinguishable: they *seem* the same, from the point of view of introspection.

According to DISJ, then, it *seeming* to me that I am having a pain experience can be one of two kinds of states of inner-awareness. But whether *it seeming to me that I am having a pain experience* is an instance of one kind of state of inner-awareness or another *makes no difference from the point of view of introspection* – in particular, it makes no *evidential* difference.

Earlier we said that the evidential grounds of phenomenal introspective beliefs are states of inner-awareness. We can see now that a disjunctivism of inner-awareness forces us to abandon this view. On this view, it is not *being a certain kind of state of inner-awareness* that is the evidential ground of phenomenal introspective belief. Rather, it is *it seeming to me that I am having a certain kind of experience* that evidentially grounds my phenomenal introspective beliefs.

And this, of course, results in phenomenal introspective beliefs *not* being conservatively indefeasible. If the evidential ground of such beliefs is common to good and bad cases, then one *can* defeat them

without changing their evidential ground. The person suffering from false pain due to dental fear will form the phenomenal introspective belief that they are in pain on the basis that it seems to them that they are having a pain experience. But this belief can be defeated, not by altering whether or not it seems to them that they are having a pain experience, but by the fact that this seeming is *not* an instance of a factive state of inner-awareness.

The SR theorist who adopts disjunctivism about inner-awareness, then, is in a similar position to the HO theorist with respect to phenomenal introspective belief. The former must hold that the common evidential factor between good and bad cases is it seeming to me that I am having a certain kind of experience, and the later must hold that the common evidential factor between good and bad cases is being in a higher-order which represents that I am having a certain kind of experience. Both views, then, cannot accommodate the conservative indefeasibility of phenomenal introspective belief.

According to Kidd, however, SR *is* in a better epistemic position with respect to such beliefs. This is because, he says:

> "there is ... no 'gap' between states of inner-awareness and corresponding first-order states, there is no place for the internal-world Cartesian skeptical worries to take root. This is not, of course, to say that one is completely free from all doubt about the epistemic adequacy of one's introspective beliefs. For indistinguishable pairs of standard and non-standard states are not banished. But one is free from the doubt that, even in a standard case of inner-awareness in experience, the experience itself still falls short of delivering grounds for knowledge." (pg. 14)

On the contrary, as I hope to have made clear, the SR theorist who adopts disjunctivism about inner-awareness *is* committed to there being a gap for internal-world skeptical worries to take root. The gap is not between states of inner-awareness and corresponding first-order states, but between one's experiences and the data of introspection: it seeming to one that one is having certain sorts of experiences.

Conclusion

Kidd attempts to demonstrate the superiority of SR theories over HO theories on the grounds that only the former can accommodate the conservative indefeasibility of phenomenal introspective belief. He accomplishes this by holding (a) that the evidential ground of such beliefs are states of inner-awareness (higher-order states that represent that one is having a certain kin of experience) and (b) that SR theorists can take such states of inner-awareness to be factive.

Kidd then attempts to render such a version of SR consistent with the fallibility of inner-awareness by adopting a disjunctive view of inner-awareness. I have attempted to show that such a disjunctivist SR theorist must reject that the evidential ground of phenomenal introspective beliefs are states of inner-awareness, and therefore cannot accommodate the conservative indefeasibility of phenomenal introspective belief. The moral: neither HO or SR theories can accommodate both the conservative indefeasiblity of phenomenal introspective belief and the fallibility of inner-awareness.