

The Regress Objection to Phenomenal Reflexive Theories of Consciousness

Daniel Stoljar, ANU

I

I will begin by reviewing three classic points in philosophy of mind.

Point 1 is that there is a theory, which I will call here ‘the reflexive theory’ or ‘RT’ for short,¹ according to which a (psychological) state² is a conscious state of a subject only if the subject of the state is conscious of being in the state. Since ‘is conscious of’ is a rough synonym of ‘is aware of’, we may also say that, according to RT, a psychological state is a conscious state of a subject only if the subject of the state is aware of being in the state.

Point 2 is that RT faces a regress objection; that is, it is apparently committed to an infinite regress of a problematic sort.³ The objection might be formulated this way. *First Premise*: if RT is true, then, if you instantiate one conscious state you instantiate an infinity of conscious states. *Second Premise*: you do not instantiate an infinity of conscious states. *Conclusion*: either RT is false or you never instantiate a conscious state—a disaster for a theory of consciousness!

The rationale behind the second premise is (apparently) straightforward: you are a finite creature, and so cannot instantiate an infinity of conscious states. Why believe the first premise? Well, suppose you are in a conscious state—call this S1. By RT, you are conscious of being in S1. But, if you are conscious of being in S1, the following three claims are (apparently) true:

The existence claim: there exists a state you are in, viz., the state of being conscious of being in S1—call this S2;

I presented versions of this paper to audiences at the Australasian Association of Philosophy Conference in Brisbane, the Australian National University Conference on Reflexive Theories of Consciousness, the University of Texas at Austin, and McGill University. I am very much indebted to the audiences at these venues, whose comments and suggestions improved the paper considerably. I am also indebted to Philippe Chuard and Declan Smithies for written comments on previous versions.

¹ Terminology: I will exclusively use the phrase ‘reflexive theory’ for the theory I have in mind, which I set out in more detail below. But the theory also goes by other names, e.g. the inner awareness thesis, the pre-reflexive theory, the self-representational theory, the self-awareness theory, and the transitivity theory. I will not go into the reasons for these different labels here.

² By ‘state’ in what follows I will always mean ‘psychological state’ unless indicated otherwise.

³ For an excellent recent discussion to which I am indebted, see Siewert 2013.

The distinctness claim: S2 is distinct from S1; and

The consciousness claim: S2 is a conscious state.

Together these three claims entail that, if S1 is a conscious state you are in, S2 is *also* conscious state you are in, and moreover S2 is distinct from S1. But then, by RT, you are conscious of being in S2, and the reasoning begins again. Hence, barring the possibility of a curious loop⁴ in consciousness, a regress follows.

Point 3 is that the proponent of RT has a standard reply to the regress objection, namely, that the consciousness claim is false, or at any rate a counterpart of that claim is false somewhere along the line.⁵ If the consciousness claim (or a counterpart of that claim) is false, there is a state with the profile of being (i) a state that consists in being conscious of something and (ii) not itself a conscious state. The simplest, though not the only, implementation of this response says that S2 in particular has that profile. If so, there is no regress and the objection is answered.

II

How plausible is this standard response to the regress objection?

In the first part of this paper, I will argue that this depends on how exactly RT is to be interpreted. RT says that *being a conscious state* and *being conscious of something* are chained together, but it does not tell us how to interpret the things chained. Moreover, an important point in contemporary philosophy of mind, which I will accept here, is that there are multiple different legitimate versions of these notions (cf. Block 1997, 2000). Hence there are multiple versions of RT, and so different potential targets for the regress objection.

In the case of (what I will call) the *cognitive version* of RT, or CRT for short, the standard response *is* plausible, or at least so I will assume here. By contrast, in the case of what I will call the *phenomenal version* of RT, or PRT for short, the standard response is *not* plausible. Hence, if one endorses the phenomenal reflexive theory, which many contemporary

⁴ I will ignore the possibility of this loop in what follows.

⁵ See, e.g., Rosenthal 2005, 9. See also Armstrong and Malcolm 1984, Gennaro 2004. Rosenthal 2011, and Lycan 2006

philosophers of mind do,⁶ some alternative response to the regress objection will have to be found.

In the second part of the paper, I go on to argue that, in the case of PRT, there is no alternative: not only is the standard response unavailable, no non-standard response is available either.

Putting these two parts together we arrive at my overall conclusion, namely, that the regress objection to the phenomenal reflexive theory is a good one, and that in consequence PRT should be rejected in favour of something else. I will end with some remarks about what this something else might be.

III

Let us begin by looking more closely at CRT and PRT, the cognitive and phenomenal versions of RT.⁷

As I will understand it, the cognitive version offers the following characterizations of the two notions at issue in RT, being a conscious state, and being conscious of something:⁸

- (1) A state T is a conscious state if and only if T is such that the subject of T knows (perhaps in a certain way) that he or she is in T; and
- (2) A subject is conscious of being in a state T if and only if the subject knows (in that way) that he or she is in T.

According to CRT, therefore, a state is a conscious state (i.e. is the object of a certain kind of knowledge) only if the person whose state it is, knows (in a certain way) that they are in that state.

The phenomenal version, by contrast, offers the following characterizations of these notions:

⁶ For philosophers who endorse this theory or something like it, see Chalmers (2010, 2013), Gertler (2010, 2012), Horgan 2012, Janzen (2008, 2011), Kriegel 2009, Nida-Rümelin 2011, and Strawson 2013. As I read them, allied strands are in: Thompson 2011, Grünbaum and Zahavi 2013 and Zahavi 2006.

⁷ As just indicated, there are multiple different versions of the notions involved in RT, and it would take us a very long time indeed to sort them all out. But it will serve our purposes here to concentrate on PRT and CRT.

⁸ The version of the cognitive reflexive theory I will work with is a version of what is sometimes called the higher order theory (HOT theory) of consciousness, but there are differences: (a) the version I discuss invokes knowledge rather than belief; (b) the version I discuss is modest in that it does not aim at giving an account of phenomenal consciousness as defined in (3) below. On modest versus immodest versions of the higher order thought theory, see Block 2011.

- (3) A state T is a conscious state if and only if T is such that there is something it is like for the subject to be in it⁹; and
- (4) A subject is conscious of being in the state T if and only if the subject is aware in a certain distinctive way—p-aware, I will say—of being in T.¹⁰

According to PRT, therefore, a state is a conscious state (i.e. there is something it is like to be in it) only if the subject who is in the state is p-aware of being in the state.

For the most part, the constitutive notions in these versions of RT are unremarkable. Both (1) and (2) appeal to knowledge—I assume we may take that idea for granted in this context. (3) is the usual definition of (phenomenal) consciousness in philosophy of mind, and, while I will say more about it immediately below, I will mostly take it for granted too. The exception here is (4), which introduces a technical idea, p-awareness. Why introduce it?

The rationale for this can be brought as follows. We have noted that ‘is conscious of’ can be interpreted in different ways, and also that it is a rough synonym of ‘is aware of’; hence ‘is aware of’ can be interpreted in different ways too. Sometimes it is used in a cognitive way to mean something close to ‘know’—that is the idea on show in (2), and indeed in CRT in general. Sometimes it is used in a perceptual way, as when we say, “Alice was aware of the sound overhead but had no idea what it was”, meaning, roughly, that Alice heard the sound overhead without knowing what it was.

Proponents of PRT, however, typically want to distance themselves from both the perceptual and the cognitive notion, at any rate as I read them. As regards the first, one may think it is possible that there is something it is like for you to have an itch even if you do not believe, and so do not know, that you have the itch.¹¹ As regards the second, one may well think that the state of having an itch (as opposed, perhaps, to the itch itself) is not something to which one can bear a perceptual relation strictly speaking. If these points are right, and if RT is true, we arrive at a key commitment of PRT as I understand it here: that there must be

⁹ It is sometimes argued that (3) by itself entails the reflexive theory or something like it. I argue against this in Stoljar 2016b.

¹⁰ It is worth noting that (4) involves what Dretske would call fact awareness or perhaps more accurately state of affairs awareness. Could one develop a version of the view that operates with object awareness? One could but I think this undermines the epistemic role of consciousness. I am very much indebted to Michael Tye for raising this point, and for help with the answer.

¹¹ This claim is controversial but I will not try to defend it here.

some relation of awareness distinct from both the cognitive and the perceptual relation; the phrase ‘p-awareness’ is a label for that relation.

Of course, that PRT is committed in this way to a technical notion of awareness raises a number of questions. Is there any such thing as p-awareness, for example, and, if so, what is its nature? These questions are important but I will set them aside. For us the key issue is not the notion of awareness with which PRT operates. It is rather the general issue of whether PRT may answer the regress objection, and in particular—this is the topic of the first part of the paper—whether the standard response described above to the regress objection is available to PRT. It is to that latter question that I now turn.

IV

The first point to make is that that the standard response *is* a good one if CRT is in play. To see this, suppose S1 is a conscious state you are in; for example, suppose it is the state of having an itch. By CRT, you know you are in S1; hence you are in the state of knowing you are in S1—call this state S2. A regress ensues only if S2 is *itself* a conscious state; could this be denied?

That is what a proponent of the standard response would say, and in the case of CRT this seems very persuasive. Surely you can know that you are having an itch and not know that you know. Indeed, that this is possible to know something without knowing that you know is something of a cliché in epistemology. It is for this reason that the standard response is plausible if CRT is in play, or at any rate so I will assume.

Turning then to PRT, to see that the standard response is *not* a good one in this case, suppose again that S1 is the state of having an itch, and that it is a conscious state. By PRT, you are p-aware of being in S1; hence you are in the state of being p-aware of being in S1—call this state S2. A regress ensues only if S2 is *itself* a conscious state; could this be denied?

That is again what a proponent of the standard response would say—but in the case of PRT that is not persuasive. For S2 meets the conditions for being a conscious state! According to the definition of a conscious state with which PRT operates—namely (3)—a state is conscious if and only if there is something it is like to be in it. But to be p-aware of having an itch *is* such that there is something it is like to be in it. In particular, what it is like to be p-aware of having an itch is *exactly the same as* what it is like have an itch: they are phenomenally equivalent. Hence by (3), the state of being conscious of S1 is itself a conscious state. So the standard response is not plausible if PRT is in play.

How might one reply to the suggestion that the standard response to the regress objection is unavailable to a proponent of PRT?

I think the most plausible move here focuses on (3), the thesis that something is a conscious state if and only if there is something it is like to be in it. I noted that this is the usual definition of phenomenal consciousness in philosophy of mind. Nevertheless, there are various ways in which this definition can be sharpened or clarified, and it is natural to think that doing so opens up a defence of the standard response.

One way to sharpen (3) is to distinguish between conscious states in what I will call the *strong* sense and conscious states in what I will call the *weak* sense, something like this:

- (5) A state T is a conscious state in the strong sense if and only if T is such that there is something it is *constitutively* like to the subject to be in T.
- (6) A state T is a conscious state in the weak sense if and only if T is such that there is something it is like to the subject to be in T, *but not constitutively so*.¹²

As an example of a conscious state in the strong sense—for short, a strong conscious state—take having an itch again. It seems reasonable to say not simply that there is something it is like to have an itch, but in addition that there is something it is *constitutively* like: you cannot have an itch without feeling in the particular way you do when you have one. As an example of a conscious state in the weak sense—for short, a weak conscious state—take the case of believing that Paris is in France. It might be that, as a matter of fact, there is something it is like to believe this; nevertheless it is possible that someone believes this and feels something different, or feels nothing at all. If so, there is something it is like to believe that Paris is in France but not constitutively so.

In the light of this, one might argue, first, that, properly understood, PRT applies exclusively to strong conscious states and so not to weak conscious states, and, second, that being p-aware of a state is only a weak conscious state. If so, a version of the standard response can be resuscitated. To illustrate, suppose again that S1 is a conscious state you are in, and suppose also that it is a strong conscious state. It then follows from PRT that you are p-aware of being in S1, and so are in S2. But it does not follow that S2 is a strong conscious

¹² In their discussion of the unity of consciousness, Bayne and Chalmers (2003/2010) distinguish a phenomenal state from a (mere) phenomenally conscious state. The distinction I describe in the text is the same as theirs (as I understand theirs), but with different terminology. The example of the belief about Paris is on loan from them.

state; for it is open to a friend of PRT to insist it is a conscious state in the weak sense but not in the strong sense. But then PRT does not apply to it, and the regress is blocked.

However, the problem here is that the claim that S2 is at most a weak conscious state is very implausible. What makes the belief that Paris is in France a weak conscious state is that, while there is in fact something it is like to be in this state, someone might have been in it and have felt nothing at all or, alternatively, something entirely different. But nothing like this is true in the case of S2, the state of being p-aware of having an itch. In particular, *anyone* who is p-aware of having an itch will feel a particular way, namely, exactly as you do when you have an itch. Hence there is something it is constitutively like to be in S2: it is a strong conscious state rather than merely a weak conscious state.

VI

Introducing (5) and (6) is one way to sharpen the definition of a conscious state offered by (3). A different way¹³ introduces the idea of a conscious state in (what I will call) the *extremely strong* sense, as follows:

- (7) A state T is a conscious state in the extremely strong sense if and only if (a) T is such that there is something it is *constitutively* like to be in it; and (b) for all phenomenally conscious states T*, T=T* if and only if what it is constitutively like to be in T* is what it is constitutively like to be in T.

As an example of a conscious state in the extremely strong sense—an extremely strong conscious state, for short—consider again having an itch. We have seen already that having an itch satisfies clause (a)—that is what makes it a strong conscious state. But off hand it also satisfies clause (b). In effect, clause (b) articulates an identity condition on phenomenal states, namely, that as far as states of this sort are concerned, phenomenal equivalence entails identity. And having an itch seems to meet that condition: on the face of it, anything that is phenomenally equivalent to an itch is an itch. Hence having an itch is not only a strong conscious state, it is also an extremely strong conscious state.

In the light of this, one might argue, first, that, properly understood, PRT applies exclusively to extremely strong conscious states and so not to either strong or to weak

¹³ I am here indebted to a suggestion by David Chalmers.

conscious states, and, second, that being p-aware of a state is not an extremely strong conscious state. If so, a version of the standard response can be resuscitated again.

Why suppose that S2 is not an extremely strong conscious state? The reasoning here is as follows. S2 is not identical to S1; indeed it is an assumption of the standard response that they are not identical. And yet they are phenomenally equivalent: what it is constitutively like to be in S2 is just what it is constitutively like to be in S1, as we have seen. Hence, S2 does not satisfy the identity condition on phenomenal states articulated in clause (b) of (7) and so is not an extremely strong conscious state.

However, the problem here is that what applies to S2 applies equally to S1. It may be that the presence of S1 bars S2 from being an extremely strong conscious state. But, *mutatis mutandis*, the presence of S2 bars S1 from being an extremely strong conscious state! After all, S1 is phenomenally equivalent to S2, but is not identical with it. So, if PRT applies only to extremely strong conscious states, it looks as if it fails to apply, not merely to S2, but to S1 as well, and that was not part of the bargain. Indeed, since the reasoning here is quite general, a proponent of PRT who endorses it is in the awkward position of having a theory that applies to no conscious states at all.

We may bring out the basic issues here in a different way by contrasting three states: having an itch, being p-aware of having an itch and believing that Paris is in France. Our original objection was that the definition of consciousness with which PRT operates entails that first two are phenomenally conscious states, in which case the standard response to the regress objection is unavailable to PRT. What we have been examining is whether we may clarify the definition to avoid this result. The outcome of the examination is that no clarification seems to work. If we work with strong conscious states, the situation is in effect unchanged: *both* having an itch and being p-aware of having an itch are conscious states, even if the belief that Paris is in France is not. And if we work with extremely conscious states, the situation is changed, but for the worse: *neither* having an itch nor being p-aware of having an itch are conscious states in this sense.

VII

The two proposals we just examined both attempt to restrict the scope of PRT by strengthening (3). There is, however, a third way we may attempt to restrict its scope, and this is to contrast conscious states in what I will call the derivative sense conscious states in what I will call the non-derivative sense.¹⁴ We may introduce these ideas as follows:

¹⁴ I am indebted to discussions the audiences at UT-Austin and McGill for raising this issue. I am particularly indebted to blah blah for help in thinking about it.

- (8) A state T is a conscious state in the derivative sense if and only if (a) T is a conscious state; and (b) there is some conscious state T* such that (i) T* is not identical to T and (ii) T is a conscious state in virtue of T*'s being a conscious state.
- (9) A state T is a conscious state in the non-derivative sense if and only if (a) T is a conscious state; and (b) T is not a conscious state in the derivative sense.

As an example of conscious state in the non-derivative sense—for short, a non-derivative conscious state—consider again our old friend having an itch. On the face of it, there is no other state in virtue of which having an itch is the conscious state it is. Hence it is conscious state and yet its being so does not derive from any other conscious state; hence, it is a non-derivative conscious state. As an example of a conscious state in the derivative sense—for short, a derivative conscious state—take the belief that Paris is the capital of France and suppose as before that it is a conscious state, i.e., there is something it is like to have it. On the face of it, this is another state in virtue of which this is state is conscious—various feelings, or images or memories for example. Hence it is a conscious state in virtue of some other conscious state; hence, it is a derivative conscious state.

In the light of this, one might argue, first, that, properly understood, PRT applies exclusively to non-derivative conscious states and so not to derivative conscious states, and, second, that being p-aware of a state is a derivative conscious state. If so, a version of the standard response can be resuscitated again.

However, there are two main problems with this proposal. The first is that it is not clear that being p-aware *is* a derivative conscious state. According to RT, a state is a conscious state of a subject only if the subject of the state is conscious of being in it. But what is the order of explanation here? On the face of it, it is natural to think that the state is conscious *because* its subject is conscious of it, and not the other way around. If so, the sstate of being p-aware is a non-derivative rather than a derivative conscious state.

Of course someone may resist this line of argument. Perhaps they will say that the order of explanation goes in the other direction. But that seems at least against the spirit of RT. An in any case, there is a second problem with the proposal. This is that it is not obvious that having an itch is a *non*-derivative conscious state; at any rate there are world-views according to which it is not. Consider the panpsychism, for example. If panpsychism is true, having an itch is a conscious state in virtue of something else being a conscious state, and so

will not itself count as non-derivative conscious state. And this was not part of the bargain. If it applies to anything, RT is supposed to apply to having an itch.

Someone might respond by denying that panpsychism is true. I don't disagree, but this doesn't matter to the present line of argument. Setting aside whether it is true, if panpsychism is possible, it is in turn possible the conscious states that we humans so not fall within the scope of PRT. That by itself that is sufficient to reject the proposal we are considering.

VIII

I have argued so far that the standard response to the regress objection is implausible if one adopts PRT. Of course it does not follow that PRT should be rejected. For it remains possible that some *non*-standard response is plausible.

What might these non-standard responses be? Well, if the standard response to the regress objection denies the consciousness claim (or a counterpart of that claim), a non-standard response would involve one of three main suggestions. First, one might deny the distinctness claim, which in this context amounts to denying that having an itch and being p-aware of having an itch are distinct states. Second, one might deny the existence claim, which in this context amounts to denying that there is a state of being p-aware of having an itch in the first place. Third, one might switch focus and deny the *second* premise of the regress objection, according to which you do not instantiate an infinite series of conscious states.

In the next few sections I will argue that none of these options is plausible either.

VIII

Suppose S1 is the state of having an itch and S2 is the state of being p-aware of having an itch. A regress ensues only if S1 is distinct from S2; that is, only if the distinctness claim is true. Maybe this could be denied? On this view, (8-10) are all true:

- (10) Having an itch = being p-aware of having an itch.
- (11) Being p-aware of having an itch = being p-aware of being p-aware of having an itch.
- (12) Being p-aware of being p-aware of having an itch = being p-aware of being p-aware of being p-aware of having an itch.

More generally, on this view, we have an infinite number of identity statements, but since the expressions involved in these statements all denote the same state, we do not have an infinite number of conscious states. This is the first non-standard reply to the regress objection I will consider.¹⁵

The problem with this view is the implausibility of these identity claims. We may bring this out by noticing that, in view of the transitivity of identity, claims such as (10-12) entail (13)

- (13) Having an itch = being p-aware of being p-aware of being p-aware of having an itch.

But (13) is, I think, implausible on its face, because the state mentioned on the right-hand side has features that the state mentioned on the left-hand side does not.

One example of such a feature is the following. In general, if you are aware of something, that thing is a potential object of attention for you. For example, if Alice is aware of the sound overhead, then the sound is a potential object of attention for her. Hence the state mentioned on the right-hand side of (13) is such that, if you are in it, the (complex) state of being p-aware of being p-aware having an itch is a potential object of attention for you. And if (13) is true, having an itch likewise has this feature. But that is implausible: mere itchiness does not enable you to attend to such a complex state.

IX

Uriah Kriegel (e.g. 2009) has defended this sort of response to the regress objection—that is, denying the distinctness claim—at some length; what would he say in response? One thing he would say is that we have confused states and properties. In particular, there is for Kriegel only one state here, S1, and it has two properties, the *qualitative* property of being such that there is something it is like to be in it, and the *subjective* property of being the object of a certain sort of awareness. But the problem with this suggestion is that Kriegel's subjective property is instantiated only if you are aware of being in S1 in a certain way—and that itself apparently involves a distinct state.

¹⁵ In fact the label 'non-standard' is a bit misleading here since it has been defended both by contemporary philosophers such as Kriegel (2009) and by historical figures such as Brentano (1979).

One might try to bolster Kriegel's view by adopting a nominalist approach to the properties question.¹⁶ On this version there is again one state here, S1, and there are two *predicates* that apply to it, the qualitative predicate 'is such that there is something it is like to be in it', and the subjective predicate 'is the object of a certain sort of awareness'. But this leaves the basic issues unaltered. If the subjective predicate is true of the state, then again it is true that you are aware of S1. And if that is true, there is again a further state you are in.

X

A different way to defend this sort of response to the regress objection draws a distinction between being conscious of something, and being conscious of something *under some appropriate mode of presentation*. We have been operating so far with (4), according to which a subject is conscious of being in a state if and only if the subject is p-aware of being in the state. But it might be suggested that (4) should be replaced with (14):

- (14) A subject is conscious of being in the state T if and only if there is some appropriate mode of presentation such that the subject is p-aware of being in T under that mode.

The interest of (14) is that one may appeal to it to defend (13) against the problem I mentioned. My criticism of (13) is that it is implausible that having an itch puts you in a position to attend to the property of being p-aware of being p-aware of having an itch. And the underlying reason for this is that it is implausible that having an itch makes you aware of such a property. But one might insist that you *are* aware of that property, and indeed can attend to it: it simply that you are not aware of it, and so cannot attend to it, under the mode of presentation suggested by the right-hand side of (13).

Now any appeal to modes of presentation raises questions beyond the scope of the present discussion.¹⁷ However, I think we can see the limitations of this defence of (13) without getting into the weeds here. For above I said that (13) is implausible because the state mentioned on its right-hand side has features that are not features of the state mentioned on its left-hand side; and I gave a fact about attention as one example of such a feature. It is true

¹⁶ I am here indebted to a suggestion by Uriah Kriegel.

¹⁷ One source of controversy is this: if modes of presentation are properties, then there is no difference between being p-aware of the relevant state under the mode of presentation *having an itch*, and being p-aware of it under the mode of presentation *being p-aware of being p-aware of having an itch*, for these properties are identical.

that the notion of a mode of presentation is relevant to this particular feature. Nevertheless, there are several *other* features here that do the job just as well. One is that the state mentioned on the right-hand side of (13) is highly complex whereas the state mentioned on its left-hand side is not highly complex. (A variant on this objection, which returns to the point that there is an infinity of true identity statements here, is that it is not the case that having an itch is an infinitely complex psychological state.) Another is that the right hand side of (13) denotes a state with (multiply) iterated intentionality, whereas the left-hand side of (13) does not. All of these suggestions I think are sufficient to reject (13). But appealing to modes of presentation makes no difference to their plausibility, for they concern the states themselves and not our capacity to attend to them.

XI

Suppose S1 is a conscious state you are in. By PRT you are conscious of being in S1. A regress ensues only if *there exists* a state you are in, S2; maybe that could be denied? On this view (15) is true but (16) is not:

(15) I am conscious of being in S1

(16) There exists a state I am in, viz., the state of being conscious of being in S1.

If that is so, the existence claim is may be denied, and the regress is blocked. This is the second non-standard response to the regress problem I will consider.

An initial problem for this view is that there is a sense of ‘state’ (the so-called ‘instantial’ sense; cf. Thau 2001) according to which (16) follows trivially from (15), just like ‘Alice and Bill are married’ entails ‘there exists a state Alice and Bill are in, viz., marriage’. If so, there is no plausibility to the denial of the existential thesis.

However, a friend of this response might draw a distinction between the instantial sense of ‘state’ and the internal sense (Thau 2001). An internal state is something located spatially in (or on) the subject (e.g. something in their brain) that is contingently identified with (or contingently identified with the realizer of) an instantial state. In sum the suggestion is (a) there is an instantial regress but no internal regress, and (b) it is only the internal regress that is objectionable.

However, there are two things wrong with this proposal. First, it is true that there a distinction between internal and instantial states, but is not as if there is no connection

between them. On some views, for example, they are in a 1-1 correspondence.¹⁸ Second, in any case, it is not true that only the internal regress is objectionable. Consider the ‘reflexive theory of liking’ according to which you like something only if you like liking it. Such a theory faces a regress. For suppose you like beans; then by the theory, you like liking beans, etc. This seems an objectionable consequence of such a theory and it doesn’t matter that the states here are one and all instantial.

XII

Suppose the first premise of the regress objection is true; suppose that if I instantiate one conscious state I instantiate an infinite number. A problem for RT ensues only if the second premise is also true, and I do not instantiate an infinite number; but maybe that premise can be denied—maybe the regress is virtuous not vicious? If so, there is no problem. That is the final non-standard response to the regress problem I want to consider.

To motivate one version of this suggestion notice that—finite creature that I am—I nevertheless instantiate an infinite number of properties. For example I have the property of weighing N kilos, and the property of having the property of weighing N kilos, and the property of having the property of having the property...etc.

However, the problem here is that regresses that involve logical properties like *instantiating* or *being true* may be virtuous; but regresses that involve psychological (and so non-logical) properties like *being conscious of something* seem not to be.

To motivate a different version of this suggestion, consider a holism about belief according to which I believe that *p* just in case (a) there is a plurality of propositions that includes *p*; and (b) I am in a state according to which that plurality is true. If the plurality of propositions is infinite, there will be an infinite series of belief reports true of me, and hence an infinite series of states—each of which is grounded in a single overall belief state. Likewise, one might suppose that there is an infinite series of p-awareness reports true of me, and hence an infinite series of states—each of which is grounded in a single conscious state.

Once again, however, there are two things wrong with this proposal. First, by itself holism does not entail that the plurality of propositions is infinite, as least as regards *ordinary* believers—in part precisely because ordinary believers are finite. Second, in cases in which it is plausible to postulate an infinite series of belief states, the belief states at issue are non-

¹⁸ A common assumption in the science of consciousness for example is that for all conscious states C, there is a neural or computational state, N, such that it is a law that subject is in C if and only if the subject is in N. The goal of the science of consciousness is to say what these laws are. See e.g. Koch 2012.

conscious or implicit. Hence there is little reason to say that the same is true in the case of conscious states.¹⁹

XIII

There are four ways to respond to the regress objection: the standard response, and three non-standard responses. I have argued that none of these responses is plausible in the case of PRT. Hence I conclude that the regress objection is a good objection against that theory, and therefore that PRT itself is false.

Even if PRT is false, however, it may be that there is some truth in its vicinity. I will end by noting two theses about phenomenal consciousness²⁰ that, while not PRT, nevertheless capture some of its spirit.

The first is to say that conscious states consist in part in a person's being conscious of something. So on this view, a (psychological) state is a conscious state of a subject only if the state consists in the subject's being conscious of something. Since 'is conscious of' is a rough synonym of 'is aware of', we may also say that, according to this theory, a psychological state is a conscious state of a subject only if the state consists in the subject's being is aware of something.

The second is to say that there is a connection between being in a phenomenally conscious state and having a belief to the effect that one is in the state at least in this sense: if one is in a conscious state, one will *believe* that one is—so long as one is rational, is interested in the issue, and is psychologically capable of forming the relevant belief.²¹

Both of these ideas capture some of the spirit of PRT. The first captures the idea that there is a connection between being a conscious state and being conscious of something. The second captures the idea that there is a connection between being in a conscious state and believing that one is—though it spells out that connection in highly dispositional and normative terms. Neither of these ideas leads to a regress, however, and it is for that reason that I recommend them here.

¹⁹ For a defence of the idea we may have infinitely iterated belief states, see Stalnaker 2009. Stalnaker makes it clear that this is most plausible when the states are non-conscious. See also Hawthorne and Magidor 2009.

²⁰ These are both theses about *phenomenal* consciousness of the sort described in (3) above. Hence they are different from CRT, which as I have understood it here is a thesis about a different (but equally legitimate) notion of consciousness.

²¹ A view of this kind is suggested by Shoemaker in various places; see, e.g., Shoemaker 2009, 1996. I have also defended a similar view in Stoljar 2016a.

References:

- Armstrong, David and Malcolm, Norman 1984: *Consciousness and Causality: A Debate on the Nature of Mind*. Oxford: Blackwell.
- Bayne, T. and Chalmers, D., 2003 'What is the unity of consciousness?' In Cleeremans 2003; reprinted in Chalmers 2010.
- Block, Ned. 1997. 'On a Confusion About a Function of Consciousness.' *In The Nature of Consciousness*, edited by N. Block, O. Flanagan, and G. Guzeldere. Cambridge, MA: MIT Press.
- Block, Ned. 2000. 'Mental Paint' In Hahn, M and Ramberg, B (eds.) 2000 *Essays in Honor of Tyler Burge* Cambridge, MA MIT Press.
- Block, N. 2011 'The Higher Order Approach to Consciousness is Defunct,' *Analysis*, Volume 71, No. 3, July 2011, 419-431.
- Brentano, F. 2015 *Psychology from an Empirical Standpoint English* London: Routledge Classics (Original Publication 1974.)
- Chalmers, D. 2010 *The Character of Consciousness* New York: Oxford University Press
- Cleeremans, A. (ed.), 2003. *The Unity of Consciousness: Binding, Integration and Dissociation*, Oxford: Oxford University Press.
- Gennaro, Rocco (ed.) 2004: *Higher Order Theories of Consciousness*. Amsterdam: John Benjamins.
- Gertler, B. 2010 *Self-Knowledge* London: Routledge
- Gertler, B. 2012 "Renewed Acquaintance" in Smithies, D and Stoljar, D. (eds) 2012 *Consciousness and Introspection* New York: Oxford University Press pp 93-128
- Grünbaum T & Zahavi, D, 2013 "Varieties of Self-Awareness" in Fulford, K.W.M et al (eds)*The Oxford Handbook of Philosophy & Psychiatry* Oxford (2013), 221-239.
- Hawthorne J and Magidor O 2009 'Assertion, Context and Epistemic Accessibility', *Mind*, 118, pp. 377-397.
- Horgan, T. 2012 'Introspection about Phenomenal Consciousness: Running the Gamut from Infallibility to Impotence.' In D. Smithies and D. Stoljar (eds.), *Introspection and Consciousness*. Oxford (2012), 405-422
- Janzen, Greg 2008: *The Reflexive Nature of Consciousness*. Philadelphia/Amsterdam: John Benjamins.
- Janzen Greg 2011: 'In Defense of the What-It-Is-Likeness of Experience' *The Southern Journal of Philosophy* 49, pp. 271-93

- Koch, Christof. 2012 *Consciousness: Confessions of a Romantic Reductionist* (Cambridge, MA: MIT Press).
- Lycan, William G. 1996: *Consciousness and Experience* MIT Press, Cambridge MA
- Kriegel, Uriah 2009: *Subjective Consciousness: A Self-representational theory*. New York: Oxford University Press
- Nida-Rümelin, Martine 2011 ‘Phenomenal Presence and Perceptual Awareness: A Subjectivist Account of Perceptual Openness to the World’ *Philosophical Issues*, 21 (1), 2011, pp. 352--383
- Rosenthal, D. 2005. *Consciousness and Mind* Oxford: Clarendon Press, 2005.
- Rosenthal, David 2011: ‘Exaggerated Reports: Reply to Block’ *Analysis* Vol 71 Number 3 July 2011 pp. 431–37
- Shoemaker, S. 2009 “Self Intimation and Second Order Belief” *Erkenntnis* 71: 35-51
- Shoemaker, Sydney. 1996. *The First Person Perspective and Other Essays*. Cambridge: Cambridge University Press.
- Siewert, C 2013 ‘Phenomenality and Self-Consciousness’ in U.Kriegel (ed) 2013 *Phenomenal Intentionality* New York: Oxford University Press, pp 235-259
- Stalnaker, R. 2009 ‘On Hawthorne and Magidor on Assertion, Context and Epistemic Accessibility’, *Mind*, vol. 118, 407, pp.399-409
- Strawson, G. 2013 ‘Self-intimation’ in *Phenomenology and the Cognitive Sciences* **12** pp 44–73.
- Stoljar, D. 2016a ‘Introspection and Necessity, *Nous*, (2016) 50 (4), pp.1-22
- Stoljar, D. 2016b ‘The Semantics of ‘What it is like’ and the Nature of Consciousness’, *Mind*, (2016) 125 (500): 1161-1198.
- Thau, M. 2001 *Consciousness and Cognition*. New York: Oxford University Press
- Thompson, E. 2011 “Memory and Reflexive Awareness”. In Mark Siderits, Evan Thompson, and Dan Zahavi (eds) *Self, no Self? Perspectives from Analytical, Phenomenological, and Indian Traditions* Oxford: Oxford University Press
- Zahavi, D. (2006) “Thinking about Self-Consciousness: Phenomenological Perspectives.” In U.Kriegel and K. Williford (eds.) *Self-Representational Approaches to Consciousness*. Cambridge MA: MIT Press

