



# Reflexive theories of consciousness and unconscious perception

Graham Peebles

Department of Philosophy, University of Geneva, Geneva, Switzerland

## ABSTRACT

A core commitment of the reflexive theory of consciousness is that conscious states are themselves necessarily the contents of mental states. The strongest argument for this claim—the necessity of inner-content for consciousness—is the argument from unconscious perception. According to this argument, we find evidence for the necessity claim from cases of alleged unconscious perception, the most well-known and widely discussed of these being blindsight. However, the reflexive theory cannot partake in this argument and therefore, must rely on at least one of the other arguments for the necessity claim. These arguments are significantly less convincing than the argument from unconscious perception, and thus the reflexive theory is left in a dialectically weak position.

## ARTICLE HISTORY

Received 14 February 2016  
Accepted 28 June 2017

## KEYWORDS

Consciousness; higher-order theories of consciousness; reflexive theories of consciousness; representationalism; unconscious perception

## 1. Introduction

I will argue that a class of representational theories of consciousness, which I will call *reflexive theories of consciousness*, cannot partake in the strongest argument for a core commitment of the view, namely the *necessity of inner-content for consciousness*. According to the family of representational theories, consciousness supervenes on some class of states with certain representational contents and functional roles. The contents of perceptual states are the things perceived—external objects, properties, and states of affairs. These are the traditional *outer-contents* of perception. In addition to these states with outer-contents, reflexive theories hold that perceptual consciousness also requires states with *inner-content*, specifically those perceptual states which represent their outer-content. Thus, in order for a state of, say, perceiving a red square to be conscious, there must not only be a state which represents the red square—that is, has this as its content—but also a state which has this state itself as its content. The necessity claim can be motivated by three arguments: the conceptual argument, the introspective argument, and

the argument from unconscious perception. The latter is based on interpretation of cases of alleged unconscious perception—cases where the subject seems to be in perceptual states but report that they undergo no conscious experience of these states—and is empirically grounded, as opposed to the other two arguments which are grounded in introspection or conceptual reflection on the notion of consciousness. It is therefore the strongest argument for the necessity claim. To editorialize somewhat, our investigation into consciousness proceeds like any other scientific investigation. Thus, any claims about consciousness which are based on, or supported by, empirical observation are on a stronger footing than those which rely either on mere conceptual reflection on what is an empirical phenomenon or on the unreliable faculty of introspection.<sup>1</sup>

For reasons I will outline below, the reflexive theory cannot partake in the argument from unconscious perception. This leaves the proponents of the reflexive theory to rest their case for the necessity claim on the conceptual and introspective arguments. The reflexive theory is thus in a dialectically weak position. I will outline some background in this section, before making the argument that the reflexive theory cannot partake in the argument from unconscious perception in Section 2. In Section 3, I will make some remarks as to the relative weakness of the other two arguments for the necessity claim.

### 1.1. Theoretical background

This paper focuses specifically on phenomenal consciousness, in particular, perceptual experience. Familiarly, states of phenomenal consciousness are states with phenomenal character, states for which there is “something it is like” to be in them: perceptual experiences, pains, pictorial imaginings, and so on. All representational theories agree on the basic core commitment that consciousness supervenes on representational states, but disagree on various further points. This paper, however, is concerned with a specific disagreement between representational theories, namely the *order of representational states and content* in the subvening base. According to first-order theories of consciousness, consciousness (as throughout, I mean non-introspective consciousness) supervenes on representational states with first-order content (of the correct type—a qualification I will omit from now on). This first-order content is outer-content—external objects, properties, and states of affairs. According to higher-order theories, consciousness supervenes on second-order representational states with the first-order state as their content.<sup>2</sup> This second-order content is inner-content—the subject’s mental states themselves. The reflexive theory is something of a hybrid of the first- and higher-order theories. According to the reflexive theory, consciousness supervenes on first-order states with both a first-order *and a reflexive second-order content*. This second-order content is reflexive because the content of the state which has it is that state itself. Consciousness thus supervenes on first-order states, as on the first-order theory, but on first- and second-order content, as on the higher-order theory.<sup>3</sup> It helps,

perhaps, to think in terms of “intentional arrows.” Let  $M$  be a first-order state and  $x$  its content. The first-order theory holds that for  $M$  to be a perceptual state in which the subject is conscious of  $x$ , it merely suffices for  $M$  to intentionally “point” (in the right way) to  $x$ . Higher-order theories hold not only that  $M$  must point to  $x$ , but also that another state  $M^*$  must point to  $M$ . So, on the first-order theory there is only one arrow that points from  $M$  to  $x$ , whereas on the higher-order theory a second arrow points from  $M^*$  to  $M$ . On the reflexive theory, this second arrow loops around from  $M$  and points back to  $M$  itself.

There is thus a significant theoretical difference between the first-order theory on the one hand, and the higher-order and reflexive theories on the other. According to the first-order theory, consciousness requires only first-order states with outer-content. According to both the reflexive and higher-order theories, for a first-order state to be conscious that first-order state itself is required to be the content of a suitably related mental state. On this view, conscious states therefore require states of inner-awareness of inner-content. The reflexive theory, however, differs from the higher-order theory in that consciousness requires conscious inner-awareness. The higher-order theory does not require the state with higher-order content to itself be conscious (otherwise the view is possibly subject to a regress, a problem that the reflexive theory does not face due to the reflexive nature of the state).

There are three arguments for the claim that inner-content is necessary for consciousness. According to the introspective argument, we find support for the view that consciousness requires inner-content from introspection. The conceptual argument, which often goes by the name of “The Transitivity Principle,” makes the case for the necessity claim by examining the concept or notion of consciousness. I will discuss these arguments in Section 3. In my opinion, the strongest argument for the necessity claim is the argument from unconscious perception. This argument is based neither on conceptual reflection on the notion of consciousness nor on introspective examination of conscious states, but on interpretation of empirical observations. The point I would like to stress in this paper is that the reflexive theory cannot partake in the argument from unconscious perception and is thus left resting on the weaker arguments for the necessity claim.

## 2. The argument from unconscious perception

There are many examples of cases of alleged unconscious perception in the literature. Some are everyday,<sup>4</sup> others are not.<sup>5</sup> The most commonly discussed non-everyday case is blindsight—indeed, blindsight is appealed to systematically in the literature as the paradigmatic empirical example of unconscious perception—and I will mostly discuss matters using the example of blindsight in this paper. Different examples are and can be cited, but their dialectical role with respect to the present discussion will always be the same.

Blindsighters appear to lack visual phenomenal consciousness in the scotoma, the affected part of their visual field. They report no phenomenal consciousness in the scotoma, however they do pass the behavioral tests, including being able to reliably report and act on objects to a degree greater than chance for vision in the scotoma.<sup>6</sup> Let us accept for the sake of argument that blindsighters pass the behavioral tests for being in a state  $M$  which satisfies the conditions for the subject's conscious visual perception of its content  $x$  as set out by the first-order theory.<sup>7</sup> If this is rejected, then the first-order theory is left untroubled by such cases. Blindsight, as with other cases of alleged unconscious perception, can be analyzed in terms of states which do not meet the required functional role for being *perceptual states*. And this is why they are not conscious. They are not conscious because they are not perceptual, not because of the lack of any inner-content. Hence, no argument for the necessity of inner-content for consciousness could be sourced from these cases. A natural interpretation, then, upon accepting that these states meet the first-order conditions for a state being conscious is that cases such as blindsight are examples of *unconscious perception*.

Recall now the intentional arrows described above. We can think of the intentional arrows as signifying the necessary, and on the higher-order and reflexive theories jointly sufficient, representational conditions, which we can call  $C$ , that representational theories set out for a state  $M$  to be a state in which its subject is conscious of its content,  $x$ . The first-order theory sets out  $C$  only in terms of the representational properties of  $M$ . The higher-order theory sets out the same conditions but also conditions for a state  $M^*$  which has the state  $M$  itself as its content. The reflexive theory's conditions involve only the state  $M$ , but  $M$  has two contents,  $x$  and  $M$  itself. The higher-order theory, then, has two subset conditions, those of the first-order theory which pertain to  $M$  and also those pertaining to  $M^*$ . The reflexive theory likewise has two subset conditions. However, the subset conditions on the reflexive theory pertain not to two different *states* but to two different *contents of the same state*. The higher-order theory splits the subset conditions by state and there is no dependence between  $M$  and  $M^*$ . Thus, the subject can be related to the content  $x$  of  $M$  in the way specified by the relevant subset of  $C$  without being related to  $x$  in the further way specified by the satisfaction of the subset conditions that pertain to  $M^*$ . The first-order conditions are the conditions for visual awareness or perception, the second-order conditions are the conditions for *conscious* visual awareness or perception. Thus, on the higher-order theory there is room for unconscious perception, something for which many think blindsight and other such examples provide empirical evidence.

As I will argue below, such a split is not possible on the reflexive theory, as opposed to the higher-order theory, as both of the subset conditions apply to the single state  $M$  and are therefore necessarily jointly satisfied if either is satisfied. Consider the respective metaphysics of the higher-order and reflexive theories. On each version of the higher-order theory, in order for  $M$  to be a state in which the subject is conscious of its content,  $M$  must be suitably related to  $M^*$  with  $M$

as its content. A question now arises for all versions of the theory. Can the subject be in  $M$  without  $M^*$ ? Although  $M^*$  may depend on  $M$ ,<sup>8</sup>  $M$  does not depend on  $M^*$ . That is, the higher-order theorist may accept that if the subject is in  $M^*$ , then they are in  $M$ ; but they do not accept that if the subject is in  $M$ , then they are in  $M^*$ . The further metaphysics of consciousness—the nature of  $M^*$  and the relation between  $M$  and  $M^*$ —are details to be worked out later. So, the basic metaphysics of the higher-order theory splits the contents  $x$  and  $M$  by state, and  $M$  and  $M^*$  are distinct states with  $x$  and  $M$  respectively as contents. The basic metaphysics of the reflexive theory, in contrast, does not split these contents by state, and  $x$  and  $M$  are both contents of  $M$ . As I will discuss presently, on any version of a *genuinely alternative* reflexive theory—and not a disguised version of the higher-order theory—the subject cannot be *partly* in state  $M$ . This is the core of the problem for the reflexive theory that I am raising. If the proponent of the reflexive theory wants to motivate the necessity of inner-content for consciousness by unconscious perception, they will be unable to offer a theory that is not a mere variant of the higher-order theory. The higher-order theory holds to two principles: the necessity of inner-content for consciousness and the independence of states with outer-content from states with inner-content. This is the way that the theory motivates the necessity claim by cases of unconscious perception. In motivating the necessity claim by cases of unconscious perception, a proponent of the reflexive theory cannot avoid also rejecting the dependence of the state with outer-content on the state with inner-content and therefore cannot avoid splitting the contents by states; which is to offer a higher-order theory and not a reflexive theory of consciousness.

We can set the argument out in terms a dilemma, stated in (2) below. The dialectic concerns how to motivate the necessity claim from cases of alleged unconscious perception, with blindsight as the chosen representative case. We begin by assuming that the blindsighter is not in a conscious visual state. The question is what moves can be made from there to motivate the necessity claim.

- (1) The blindsighter is not in a conscious visual state.
- (2) Either the blindsighter is or is not in a state which meets the first-order theory's conditions for visual consciousness.
- (3) If they are not, then they are not in a state of unconscious perception—and, by extension, there are no cases of unconscious perception as all such cases can be handled in the same manner by the first-order theory which rejects unconscious perception.
- (4) If there is no unconscious perception, unconscious perception cannot motivate the *necessity claim* that a further state which has the first-order state as its content is required for the first-order state to be conscious.
- (5) In order to motivate the necessity claim, the blindsighter needs to be in a state which meets the first-order theory's conditions for visual consciousness.

- (6) The blindsighter cannot be in any further state which renders the visual state conscious.
- (7) Therefore, the state that the blindsighter is in must be able to *obtain independently* of any such state.
- (8) The conjunction of the *necessity claim* and this *independence claim* defines the higher-order theory.
- (9) Therefore, unconscious perception cannot motivate any theory other than the higher-order theory.

(1) is an assumption. (2) is an application of the law of excluded middle. (3) follows from the acceptance that if they are not in a state which meets the first-order theory's *C*, the second disjunct in (2), then they are not in a perceptual state—and if one accepts this for blindsight then one presumably accepts that all alleged cases of unconscious perception can be handled by the proponent of the first-order theory in this way. If blindsight and other cases of alleged unconscious perception can be handled by the first-order theory through its denial of unconscious *perceptual* states, then neither the reflexive nor the higher-order theory can source an argument for the necessity claim on the basis of these cases. Hence (4). The second disjunct in (2) is a dead end for the reflexive theory. The proponent of the necessity claim must therefore accept the first disjunct in (2). That is, in order to motivate the necessity claim on the basis of cases of unconscious perception, one must hold that the blindsighter is in a state which meets the first-order theory's *C*—(5). (6) is clear and follows from (1). So (7) follows. (8) is a definition. Hence, we conclude (9). I will defend steps (7) and (8) of this argument in the rest of Section 2.

### 2.1. *The metaphysics of the higher-order and reflexive theories*

The feature of the reflexive theory that I am arguing deprives it of the explanatory advantage that the higher-order theory has in explaining unconscious perception does not rest on the precise metaphysics of the constitution of conscious states but on the *dependence* between the outer and inner-content (7). As the two are contents of the same state, outer-content must come with inner-content. If not—if the subject could be aware of the content *x*, the outer-content, without being aware of *M*, the inner-content—then the theory that accommodates this is not the reflexive theory but some variant of the higher-order theory (8). This is the core metaphysical difference between the two views in respect of the constitution of conscious states. The higher-order theory splits the contents by states, whereas the reflexive theory does not. If a theory holds to the independence of the state with outer-content from the state with inner-content, then it is a higher-order theory and not a reflexive one: independence goes hand in hand with splitting the contents by state. This is again to state (8), which combined with (7) gets us to (9).

The point being made here sidesteps the specifics of the further metaphysics of both the reflexive and higher-order theories. Just as the higher-order theory

comes in different versions, so does the reflexive theory. The two main options are that the two elements of the state with the inner- and outer-contents are parts of the same whole, or that the states with the different contents are identical.<sup>9</sup> These further metaphysical questions, though, are not really pertinent to the matter at hand. As noted, the argument being made here turns on the dependence of the state with outer-content on the state with inner-content.

I am arguing that any theory which abjures the independence of outer-content from inner-content cannot partake in the argument from unconscious perception. But if a theory accepts this independence, then in what way is such a theory offering an alternative to the higher-order theory? What does it mean to say that such theories form a genuinely different class of representational theory? According to the general higher-order view,  $M$  and  $M^*$  are distinct states where  $M$  does not depend on  $M^*$ . To present a view, where  $M$  and  $M^*$  are identical but  $M$  does not depend on  $M^*$  seems to me to be incoherent. How can  $M$  and  $M^*$  be identical and yet one not depend on the other for its obtaining? Further, is to say that  $M$  and  $M^*$  are distinct parts of a complex state  $M/M^*$  where either there is no dependence between the two or  $M^*$  depends on  $M$  but not vice versa, really to say something different from the higher-order theory in one of its variations? It seems to me not. It seems to me that such a view is but a mere variation on the higher-order theory. If there is a genuine difference between the reflexive and higher-order theories with respect to the metaphysical constitution of  $M/M^*$ , then it turns on the acceptance of the dependence of  $M$  on  $M^*$ , regardless of the metaphysics, whatever this may be, of the complex state  $M/M^*$ .<sup>10</sup>

Consider, then, the following way of individuating representational theories of consciousness according to whether or not they accept the necessity of inner-content for consciousness (necessity) and/or the dependence of states with outer-content on states with inner-content (dependence).<sup>11</sup>

Dependence	Necessity	Theory
no	no	first-order
no	yes	higher-order
yes	yes	reflexive

First-order theories do not require that a state with outer-content itself be the content of a mental state in order for it to be conscious, and the state with outer-content does not depend in any way on another state. The first-order state meets the conditions for perception, which are also the conditions for conscious perception on this theory. Reflexive and higher-order theories, on the other hand, hold that for a state with outer-content to be conscious, it must itself be the content of a mental state. The fundamental difference between the higher-order and what I am calling reflexive theories turns, then, on dependence. On the higher-order theory, dependence is rejected and a state  $M$  can meet the conditions  $C$  set out by the first-order theory without that state being conscious. Any theory, however, which affirms that the state  $M$ —the state which first-order theories take to meet

their  $C$ —depends on the state  $M^*$  with  $M$  as its content for its obtaining, in addition to the thesis that  $M^*$  and  $M$  are individually necessary and jointly sufficient for  $M$  to be conscious, cannot appeal to blindsight and other such cases of alleged unconscious perception as support for the view that inner-content is necessary for consciousness. On such a view, if a subject is in  $M$  (is in a state which meets the first-order  $C$ ) then as this state depends on  $M^*$ ,  $M$  would necessarily be a *conscious state*. If, however, the theory allows for the blindsighter, for example, to be in  $M$  but not in  $M^*$ , then the view on offer is a version of the higher-order theory. This is (8) as set out above. And (7)—that if the blindsighter is in a state which meets the first-order  $C$  but is not visually conscious, then this state must be able to obtain independently of a further state which would render it conscious—makes clear the dialectical difficulty that the proponent of the reflexive theory is in with respect to cases of unconscious perception. (1) is an assumption that is required to get the debate off the ground. (2) is an application of excluded middle, and poses a dilemma. The reflexive theorist cannot hold that the blindsighter does not meet the first-order theory's  $C$ , for then the first-order theory can deal with these cases without recourse to the necessity of inner-content for consciousness. So, the reflexive theorist must hold that the blindsighter meets the first-order  $C$ . However, as per (1), the blindsighter is not visually conscious and so, (6), cannot be in any further state which would render that state conscious. Hence, their visual state must be *independent* of any state which would render it conscious (7). But now we have the conjunction of two claims, necessity and the rejection of dependence. Necessity is the thesis that inner-content is required for consciousness. To reject dependence is to reject the claim that perceptual states depend on the states with inner-content which render them conscious for their obtaining. But this conjunction is the higher-order theory. The proponent of the reflexive theory cannot avoid rejecting dependence if they base their argument for necessity on cases of unconscious perception.

This, then, poses a question to the proponent of the reflexive theory, no matter the metaphysics of conscious states—for example, part/whole or identity—on any particular version of the view. We can group many theories together under the general umbrella of “higher-order theory.” The dispositionalist, actualist, perceptual, and thought versions of the view all hold to the acceptance of necessity but the rejection of dependence. If a theory also holds to this, as it must if it is motivated by unconscious perception, then in what sense is it an *alternative* to the general higher-order view as opposed to some version of it? The only theory that unconscious perception can motivate is the higher-order theory. Of course, the labels “higher-order” and “reflexive” are just that—labels. But the underlying point remains. All higher-order theories allow for the interpretation of, for example, blindsight that the subject perceives but is not conscious because they are not in a state with inner-content. They are perceptually aware of the content of their perceptual state but not consciously so. This, so the explanation goes, is due to the lack of inner-content—due to the non-obtaining of an independent state

with inner-content. The argument from unconscious perception is the strongest argument for the necessity claim, and if the higher-order theory alone can lay claim to it, where does that leave the reflexive theory? Either one rejects the unconscious perception interpretation of cases such as blindsight, and therefore, essentially sides with the first-order theorist in this argument, or one accepts the interpretation and sides with the higher-order theorist.

## 2.2. A substantive problem?

Matters, however, may appear to be more complicated.<sup>12</sup> A look at a feature of the theory in Gennaro (2012) perhaps helps to draw this out. According to Gennaro's theory, first-order conscious states are conscious because they contain *intrinsic* higher-order thoughts. This theory is designed to steer something of a middle course between explicitly higher-order theories, such as Rosenthal's, and explicitly reflexive theories such as Kriegel's.<sup>13</sup> According to a view such as Rosenthal's, the state  $M$  is independent of the state  $M^*$ . When  $M$  becomes related to  $M^*$  in the correct way,  $M$  becomes conscious and the subject is consciously aware of its content. Crucially,  $M$  can occur without  $M^*$ . Gennaro's view is subtly different from standard higher-order theories with respect to the interaction between  $M$  and  $M^*$  and the resulting complex of these elements. According to Gennaro's theory:

What makes mental states conscious is intrinsic to conscious states, but a kind of inner self-referential and relational element is also present within the structure of such states. In contrast to standard [higher-order] theory, [Gennaro's theory] says that first-order conscious mental states are complex states containing both a world-directed mental state-part  $M$  and an unconscious metapsychological thought ... It is, if you will, an intrinsic version of [the higher-order thought] theory (Gennaro, 2012, p. 55).

The metaphysical dispute between, say, Rosenthal—here representing higher-order theories of all kinds—and Gennaro concerns the metaphysics of conscious states involving two elements. According to Rosenthal, these elements are distinct states, whereas according to Gennaro they are both intrinsic elements of the same state. Let us consider a case where this first-order element occurs without the further element, followed by the occurrence of the further element which will render the subject conscious of the content of the prior state.

The difference between the two views is that *the same state  $M$  becomes* conscious (or at least a state of the same type—see below) when targeted by  $M^*$  on Rosenthal's view, whereas this state is *replaced* by a conscious state on Gennaro's view. Why, then, could a proponent of the reflexive theory not respond in a similar way? On this view, the subject could be in a state  $M$  with content  $x$ . When the subject moves into state  $M^*$  which has contents  $x$  and  $M$  they come to be in a conscious state because  $M^*$  is *itself conscious* and not that  $M^*$  *makes  $M$  conscious*. Rather, as with Gennaro's view,  $M$  is replaced by  $M^*$ . It may, then, appear that such a move can explain the difference between conscious and unconscious states: the blindsighter, for example, is in an unconscious perceptual state, which

would be replaced by a conscious perceptual state which has a different structure (i.e., content and functional role) were they to become conscious of the content of the prior state. The reflexive theorist may argue that this allows them to hold that the blindsighter is in a state  $M$  with outer-content which meets the first-order  $C$ .  $M$  is then replaced by  $M^*$  which duplicates  $M$  as a part but has an extra layer of reflexive content.  $M$  is not conscious but  $M^*$  is. This maneuver, however, will not work.

Consider again how the reflexive theorist could respond to blindsight cases and recall the dilemma set out above. Either the blindsighter is or is not in a state which meets the first-order theory's  $C$  for being in a state of visual consciousness (2). If they are not, then there is no leverage here for either the proponents of the reflexive or higher-order theories to motivate the introduction of inner-content as a necessary condition for visual consciousness (3/4). So in order to motivate the necessity claim, the blindsighter must be understood as being in a state which in fact meets the first-order  $C$  despite not being in a conscious state (5). Thus, space is made to introduce inner-content as a further condition on the state being conscious. The dialectic then moves on to consider what these cases of unconscious perception can show us about which theory which includes this necessity claim is the correct one. The blindsighter is not visually conscious and so cannot be in any state which would render the visual state conscious (6). Hence, dependence must get rejected (7). First-order perceptual states do not depend on states with these states themselves as their contents. And this is the higher-order theory, not the reflexive theory (8).

The reason why the response now being considered will not work, and why there is no escape from (7) and (8), concerns the way that states are individuated on the representational theory: dependence is a question of state *types*. We need not look to such exciting cases as blindsight to draw this out. Consider, for example, Armstrong's sleepy lorry driver. He drives along unaware, according to this interpretation, of the state which obviously guides his action in exactly the same way with respect to the input of perceptual stimuli and the output of action as a corresponding conscious state would, only this state is not conscious. At some point in the journey, he becomes conscious. Whatever causes this, we need not assume any difference either in the input or the behavioral output; at some point, and for some reason, he becomes conscious with neither the content of his perceptual state nor its action-guiding functional role changing—otherwise, the first-order theorist will step in to challenge the prior state's meeting their  $C$ . The higher-order and reflexive theorist will, of course, hold that there is a difference in the *overall* functional role. But if they argue that the prior states meet the first-order  $C$ , as they must to ward off the first-order theory, then they are surely forced to hold that the functional role with respect to the *strictly perceptual* nature of the state is retained by the latter state. So, it is difficult for me to see how a case can be made that the former state is replaced by a state all of whose parts are type-distinct from the previous state and cannot obtain without the other parts. For this has

been explicitly denied at the off, when it is argued that the prior state meets the first-order theory's *C*.

We need not embroil ourselves in complicated metaphysics here to see the point. If a subject is in a state *M* which meets the first-order *C* and then comes to be in a state *M\** which duplicates state *M* as part of its overall functional-representational properties, then surely the part of *M\** with this content and functional role is a state of *the same type* as *M* which, as on the higher-order theory, does not depend on whatever else constitutes the rest of *M\** for its obtaining. We can set aside questions of token identity here. Whichever way one goes on the question of token identity of mental states over time, our question concerns whether tokens of the type with content and the functional role that the first-order theory sets out for perceptual states can obtain independently of (relevant) states which have these states as their content. If they can, then dependence is rejected.<sup>14</sup>

So, consider a final time the unconscious-perception interpretation of Armstrong's sleep lorry driver. The debate is internecine between representationalists.<sup>15</sup> All parties accept at least minimal representationalism. Consciousness supervenes on representational content and functional role, and it is to these properties that we look when identifying not only why states are conscious or unconscious or what makes them conscious or unconscious, but what type a state is: is it a perceptual state or a cognitive state, and so on? The proponent of the necessity claim must hold that he is in a state which meets the first-order theorist's *C* in order to ward off the first-order theory. This view is reached by considering the representational-functional properties of the state as judged by its content and behavioral output. This is, so to speak, the game that we are playing: the representationalist game of individuating state types. Now, when driver "comes to," he is in a state which has the same behavioral output with respect to navigating the road and therefore the same functional-representational properties responsible for this behavioral output as the previous state. Of course, the overall content and functional role will be different—he is now in a conscious visual state—but the overall content and functional role contains the previous content and functional role as a subset. Any state that the driver was in as he drove along the road could have been a conscious or an unconscious state—but it could not have had a different content or functional role with respect to its mediating input stimuli and output behavior.<sup>16</sup> For this defines what type of state it is on a representational theory, and in order to secure the necessity claim the states prior to his becoming conscious must be held to be perceptual. Thus, the higher-order theory's response is inescapable: perceptual states require inner-content to be conscious but do not depend on states with inner-content for their obtaining; therefore, the states with perceptual outer-content are distinct states from those with inner-content. If the proponent of a *reflexive* theory, one which holds that the two contents are contents of the same state, is to motivate the necessity claim, they will therefore have to look elsewhere.

### 3. The conceptual and introspective arguments

It is not my goal in this paper to attempt to refute the reflexive theory or the necessity claim, but rather to make the dialectical point that the reflexive theory has greater difficulty in motivating the necessity claim than the higher-order theory due to its inability to appeal to the argument from unconscious perception and its resultant reliance on either the introspective or the conceptual argument. In this section, I will briefly survey the objections to these arguments on their own terms. In my view, they are methodologically weaker than the argument from unconscious perception due to the fact that the argument from unconscious perception alone rests on an empirical basis, but I cannot argue for this in depth here and I therefore leave this matter up to the reader.

#### 3.1. The argument from introspection

The argument from introspection is that we introspect our inner-awareness of our conscious states. However, appeal to introspection for the necessity claim will not succeed. All states of introspection are, by definition, states of inner-awareness with inner-content. Of course we are aware of our first-order awareness when we introspect, and are therefore in a state with inner-content. A simple appeal to introspection, therefore, for the claim that self-awareness is a necessary requirement to be in a conscious state can be dismissed. It is not necessary, but merely sufficient. Kriegel (2009b), however, has a stronger version of the introspective argument. Consider, for example, a small scratch right in the center of a computer screen that one is in some sense conscious of—it is right in the middle of the visual field—but in some sense “tunes out.” According to Kriegel, we are always at the least *peripherally* aware of our awareness of such things. This peripheral awareness is the analysis of the “tuning out.” When we shift our attention to this awareness, of which we are only initially peripherally aware, by introspecting it, we become introspectively aware of this awareness. According to Kriegel, what happens here is that introspection supplants peripheral awareness. One was peripherally aware of some awareness (and focally aware of the rest), and one then becomes introspectively aware of this awareness of which one was hitherto only non-introspectively peripherally aware. At all times, then, the outer-awareness is the inner-content of one of two forms of inner-awareness, either introspective or peripheral. But why should we believe in peripheral awareness? The first-order theorist—who rejects the necessity of inner-content—will argue that the shift from a non-introspected first-order state to an introspected first-order state is just that, an act of introspection. That one is in a state with inner-content when one introspects can only show something about introspection, not the non-introspective state which preceded it.

However, to be precise, the argument here is that we are non-introspectively—that is, peripherally—aware of our outer-awareness at all times; as noted, introspection displaces peripheral awareness. Thus, the claim that consciousness

always comes with a state with inner-content is not motivated by introspection, but by merely noticing one's phenomenology. Self-awareness is something that one notices in essentially the same way as one notices external objects and their properties. The reflexive theory is alone in being able to appeal to this argument, as higher-order theories hold that the higher-order states which are required for the lower-order states which are their contents to be conscious are themselves unconscious. These theories thus cannot appeal to phenomenology for the necessity claim. However, even understood in this way—where introspection displaces peripheral awareness—the argument does not succeed. There are many objections to this notion of peripheral awareness, but I will not discuss them in detail here.<sup>17</sup> I will make only two points against the notion, the first being more of a statement than an argument. It is in my view just untrue as a matter of psychology (statement), and, more speculatively, it is really the conceptual argument in disguise.

Any claim which rests on introspection falls to the rebuttal that it can merely establish sufficiency not necessity. Kriegel's claim—that we are all non-introspectively in a peripheral state with inner-content—does not fall to this objection. In my opinion, however, it does fall to the objection that it is simply untrue as a matter of psychology. If such a thing is indeed at all possible, when I try to catch myself taking in my conscious states—not taking in the *first-order content* of my conscious state but those states themselves—*without introspecting*, I find no such peripheral awareness. Indeed, I would perhaps venture to go so far as to say that the claim that there is always such peripheral awareness is quite obviously untrue. However, others will affirm the contrary with equal conviction, and it is not clear how much progress can be made with this dispute on these terms.

There is, though, also a further problem for the argument which does not result from any such psychological claim. Even if one were always to find peripheral inner-awareness accompanying conscious states, this would only show that *as a matter of fact* such peripheral inner-awareness always accompanies conscious states, not that this is *necessarily so*. In order to show this, one would have to adduce a stronger argument. The only remaining argument that the proponent of the reflexive theory can appeal to here is the conceptual argument.

### 3.2. The conceptual argument

In a paradigm example of this argument, Lycan offers a “simple argument” for the view that “a conscious state is a state that is itself represented by another of the subject's mental states” (Lycan, 2001, p. 3). He begins his argument for this with the “stipulative definition” that “a conscious state is a mental state whose subject is aware of being in.” As has been pointed out, however, without independent motivation this claim will not cut much ice.<sup>18</sup> But what motivation could there be for this? Setting aside the possible empirical and introspective motivations, we are left with the conceptual. Rosenthal makes clear what I think is the essence of the conceptual argument.

[Inner-content] must in any case occur if there is something it's like to have the experience. We're not interested in there being something it is like for somebody else to have the experience; there must be something it is like for one to have it, oneself. Without specifying that, what it's like would be on a par with what it's like to be a table. (Rosenthal, 2002, p. 656)<sup>19</sup>

However, nothing said here implies anything about inner-awareness, inner-content, or any such notion. But something like this likely stands behind Lycan's motivation for his first premise. Unfortunately for the proponent of the necessity claim, however, we can make perfect sense of all this with the resources of the first-order theory. There is no valid transition from an experience *being mine* to my *being aware* of that experience, let alone the claim that this awareness is conscious. All that we need say is just that, that the experience is mine. The conditions *C* that the first-order theory sets out are satisfied by a state of one subject as opposed to another. The state is a state of *this* subject as opposed to a state of *that* subject. The qualification "for the subject" marks mere ownership of the state. Nothing in the fact that one state is a state of, or for, one person and another of, or for, another person pushes us to say any more than this. Compare this now to the way in which Kriegel introduces his twofold characterization of phenomenal character on the first page of *Subjective Consciousness*:

When I have a conscious experience of the blue sky, there is something it is like for me to have the experience. In particular, there is a bluish way it is like for me to have it. This "bluish way it is like for me" constitutes the phenomenal character of my experience. Phenomenal character is the property that makes a phenomenally conscious state (i) the phenomenally conscious state it is and (ii) a phenomenally conscious state at all. The bluish way it is like for me has two distinguishable components: (i) the bluish component and (ii) the for-me component. I call the former qualitative character and the latter subjective character ... phenomenal character is just the compresence of qualitative character and subjective character ... [and] ... a phenomenally conscious state's qualitative character is what makes it the phenomenally conscious state it is, while its subjective character is what makes it a phenomenally conscious state at all. Thus, my conscious experience of the blue sky is the conscious experience it is in virtue of its bluishness, but it is a conscious experience at all in virtue of its for-me-ness.

Here we can see the conceptual argument—note how Kriegel distinguishes the mere "bluish way it is like" and the bluish way it is like "for me"—and the introspective/peripheral-awareness argument both intertwined. However, as both arguments are independently questionable, their conjunction will not strengthen either. We do not have peripheral awareness of inner-content, and even if we did this would not establish the necessity claim. The only way to establish the necessity claim would be by the conceptual argument, but the conceptual argument is flawed.

#### 4. Conclusion

Constraints on length preclude any extensive discussion of these points. My central aim in this paper was to argue that the reflexive theory cannot partake in the

argument from unconscious perception for the claim that inner-content is necessary for consciousness, and that the view is thus left with either the introspective or conceptual arguments as motivation. In my view, both of these arguments are independently weaker than the argument from unconscious perception—both methodologically and on their own terms. I noted above that I cannot defend the methodological principle that empirically motivated arguments, such as those from observation of cases of alleged unconscious perception, carry more weight than conceptual or a priori arguments, and I must therefore hope that the reader shares this view. All that I can say to this here is that the approach to consciousness under which the debates discussed in this paper proceed is in large part empirically minded. Not all in the debate, of course, would agree with this. And one should not read this as implying that the best method is solely empirical and anti-philosophical. On the contrary, the representationalist program is one which has sought to bring together the empirical and philosophical literature. Hence, in this debate, I would say that empirically based, or at the very least informed, arguments carry more weight than merely conceptual or introspective arguments, a representative example of such an empirically based argument being the argument from unconscious perception for the higher-order representational theory.<sup>20</sup>

Both the arguments from introspection and the conceptual argument are on their own terms weaker than the argument from unconscious perception. The greatest strength of this argument is the difficult position it places the proponent of the first-order theory in with regard to explaining the actions of blindsighters and subjects in similar states, and it is thus a strong argument for the necessity claim. The argument from introspection does not succeed. The conceptual argument is likewise not convincing. The relation between a state's being conscious and that state being a state of its subject, or a state that the subject is aware of being in, can be understood with the resources of the first-order theory, resources which do not include the necessity of inner-content for consciousness. Hence, as the reflexive theory cannot partake in the argument from unconscious perception, it is in a dialectically weak position.

## Notes

1. See Schwitzgebel (2008) for discussion.
2. Notable proponents of variants of the respective theories include: Harman (1990), Tye (2000), Dretske (1995), and Armstrong (1968), Lycan (1996), Rosenthal (2005), Carruthers (2000), (2005), Gennaro (2012). Higher-order theories come in two different kinds. Actualist theories require the first-order state to actually be the content of a higher-order state, whereas dispositionalist theories require only that it be disposed to be so. Another distinction between higher-order theories is that between those which hold that the higher-order state is cognitive and those which hold that it is perception-like. This difference, as with the other differences between representational views not concerned with the order of states and contents, is not relevant to this paper.

3. For various versions of the view, see Kriegel (2009a, 2009b), a number of the papers in Kriegel and Williford (2006), Hossack (2002), Janzen (2008), Caston (2002), and Textor (2013a, 2013b).
4. A notable such example is Armstrong's (1981) sleepy lorry driver, who suddenly "comes to" while driving and seems to realize that he previously was unconscious of the road in front of him. Nevertheless, he must have perceived it in order to successfully have driven it.
5. See Lau and Rosenthal (2011) for an overview.
6. See Weiskrantz (1997), Vision (1998), and Holt (1999) for general overviews, and Carruthers (2000) for philosophical discussion.
7. Many find the arguments for this compelling, although they are to my mind far from conclusive. The proponent of the first-order theory can argue, for example, that the functional role of the state which explains their success in acting and reporting on the contents in the scotoma does not *quite* satisfy the conditions set out by the first-order theory for visual consciousness (For example, Tye, 2000, pp. 62–63). Others are skeptical of this response (see, for example, Carruthers, 2000, Chapter 6 for critical discussion of its merits), and hold that blindsight demonstrates that the conditions set out in the first-order theory are not sufficient for perceptual consciousness, but merely sufficient for visual awareness. Blindsighters, so this line of reasoning goes, must be perceptually aware of the content in the scotoma as they are able to reliably act and report on it. As they demonstrate sufficient similarity in abilities to act and report on the contents of the scotoma to normally sighted subjects, those who object to the first-order theory argue that the subjects' mental states which explain this therefore do meet the conditions set out on the first-order theory. The test for this is sufficient behavioral similarity to normally sighted subjects, tests which blindsighters pass. Hence, something must be added to the first-order theory to render it a sufficient account of perceptual consciousness. The proponent of the higher-order theory concludes, again, that what is missing is a suitable account of inner-content. The blindsighter is aware of the content of the perceptual state, as shown by observation, but unaware of that awareness itself. And so, the perceptual awareness is unconscious. These are still matters of dispute in the literature, and I will make no attempt to settle the question here. My argument concerns only the effectiveness of the argument from unconscious perception for the reflexive theory.
8. This is a matter of dispute, and is indeed something of a problem for the view, as some argue that the theory allows for "targetless" higher-order states which do not have as their contents actual lower-order states (for example, Block, 2011; Byrne, 1997; Neander, 1998). It would then seem to one as though one was in a conscious first-order state without actually being in one. The higher and lower-order states may also be misaligned in that that content of the  $M^*$  that the subject is in may not be  $M$  but another mental state  $N$  with a different content, even though the subject is in  $M$ , not  $N$ . For the purposes of this paper, however, we can set these discussions aside. What matters here is the dependence of outer-content on inner-content, not the converse.
9. See Kriegel (2006) for discussion.
10. The reflexive theory differs in that it holds that inner-awareness is conscious awareness. However, this cannot be established by the argument from unconscious perception and must be established by a different argument (see Section 3). This part of the paper concerns the argument from unconscious perception and only what it can tell us about the necessity claim, which does not include the claim of conscious inner-awareness.

11. This table omits the fourth option. It is not clear if the fourth option is feasible. As the view rejects necessity, it would have to be a version of the first-order theory, but why then accept dependence?
12. I would like to thank anonymous referees for pressing for some more detail here.
13. See Gennaro (2012, chapter 5) for comparison between his view and reflexive theories.
14. The argument therefore does not rely on a principle along the following lines to block the reflexive theorist's claim to cases of unconscious perception: if the functional-representational properties of state  $M$  and the content  $x$  are subsets of the functional-representational properties of state  $M^*$  and content  $x^*$ , then if state  $M$  with content  $x$  obtains then so does state  $M^*$  with content  $x^*$ . This would make unconscious perception impossible for the reflexive theory, but it is false as a general principle. Rather, the argument is that the only way to allow  $M$  to obtain without  $M^*$  and therefore for  $M$  be a state of unconscious perception which could motivate the necessity claim is to hold that  $M$  is independent from the state  $M^*$ . And this is to offer a higher-order theory of consciousness where the state with the content that the subject is conscious of is a distinct state from that state which takes this prior state as its content and therefore renders it conscious. This is opposed to the reflexive theory which takes the two contents to be contents of the same state.
15. Or intentionalists, if one prefers.
16. Strictly speaking, it could differ in that the actual output is a subset of some possible outputs—the driver could have had better eyesight or night vision, for example. But it could not have different output to a degree which changes its functional role enough to allow the first-order theorist to challenge its meeting their  $C$ .
17. See Gennaro (2012, chapter 5) for comprehensive discussion.
18. See Lurz (2001).
19. This particular passage was brought to my attention in Byrne (2006), who argues against the higher-order theory by questioning the necessity claim and arguing that the first-order theory's resources are all that is required. Proponents of higher-order and reflexive theories differ in the weight that they place on the Transitivity Principle, that a state is conscious when the subject of the state is conscious of that state (Rosenthal, 1997), and the role that it plays in their theory. Some—Lycan, for example—offer it as an argument for their theory. However, others—Rosenthal, for example—place less argumentative weight on it, offering it as a truth about consciousness which allows us to zero-in on exactly what consciousness is, prior to offering a theory of consciousness or an analysis of the notion. Nevertheless, the following seems to me a reasonable question. Whatever philosophers' intentions are when it comes to the Transitivity Principle in this respect, if one is guided in one's investigation into consciousness by the Transitivity Principle, then if no other argument is offered for the necessity of inner-content for consciousness, in what way is the Transitivity Principle not at least implicitly playing the role of an argument for the necessity claim? One may resist the charge that what is being offered is an argument rather than an observation about consciousness which allows us to properly understand the phenomenon under investigation, but my response to this would be to ask how what is taken to be an observation about consciousness, and therefore a truth about it, could not be playing a role in one's *theory* of consciousness. Consciousness is an extremely tricky notion to zero-in on, and the Transitivity Principle, even if not intended as part of a theory of consciousness, does at least place restrictions on what that theory can in the end be if it is understood to be a substantial truth about consciousness: if the Transitivity Principle is correct, then this theory must reflect the relation between a state's being conscious and the subject's awareness of that state, where this is understood in a

non-trivial way. I lack the space here to elaborate on the role that it plays or could play, implicitly or explicitly, in the arguments of individual philosophers, and so I will confine myself to the following remarks on the Transitivity Principle and the conceptual argument as I believe that they broadly function in the general debate.

20. See, for example, Lau and Rosenthal (2011).

## Acknowledgment

I would like to thank Li Zhongwei for discussion of the contents of this paper, anonymous referees for their comments, and Akiko Frischhut for comments on drafts of this paper.

## Disclosure statement

No potential conflict of interest was reported by the author.

## References

- Armstrong, D. M. (1981). What is consciousness? In D. Armstrong (Ed.), *The nature of mind and other essays* (pp. 721–728). Ithaca, NY: Cornell University Press.
- Armstrong, D. M. (1968). *A materialist theory of the mind*. London: Routledge.
- Block, N. (2011). The higher order approach to consciousness is defunct. *Analysis*, 71, 419–431.
- Byrne, A. (2006). What phenomenal consciousness is like. In R. Gennaro (Ed.), *Higher-order theories of consciousness* (pp. 203–225). Philadelphia, PA: John Benjamins.
- Byrne, A. (1997). Some like it HOT: Consciousness and higher-order thoughts. *Philosophical Studies*, 86, 103–129.
- Carruthers, P. (2005). *Consciousness: Essays from a higher-order perspective*. Oxford: Oxford University Press.
- Carruthers, P. (2000). *Phenomenal consciousness*. Cambridge: Cambridge University Press.
- Caston, V. (2002). Aristotle on consciousness. *Mind*, 111, 751–815.
- Dretske, F. (1995). *Naturalizing the mind*. Cambridge, MA: MIT Press.
- Gennaro, R. (2012). *The consciousness paradox*. Cambridge, MA: MIT Press.
- Harman, G. (1990). The intrinsic quality of experience. *Philosophical Perspectives*, 4, 31–52.
- Holt, J. (1999). Blindsight in debates about qualia. *Journal of Consciousness Studies*, 6, 54–71.
- Hossack, K. (2002). Self-knowledge and consciousness. *Proceedings of the Aristotelian Society*, 102, 163–181.
- Janzen, G. (2008). *The reflexive nature of consciousness*. Amsterdam: John Benjamins.
- Kriegel, U. (2006). The same-order monitoring theory of consciousness. In U. Kriegel & K. Williford (Eds.), *Self-presentational approaches to consciousness* (pp. q43–170). Cambridge, MA: MIT Press.
- Kriegel, U. (2009a). *Subjective consciousness: A self-presentational theory*. Oxford: Oxford University Press.
- Kriegel, U. (2009b). Self-representationalism and phenomenology. *Philosophical Studies*, 143, 357–381.
- Kriegel, U., & Williford, K. (Eds.). (2006). *Self-presentational approaches to consciousness*. Cambridge, MA: MIT Press.
- Lau, H., & Rosenthal, D. (2011). Empirical support for higher-order theories of conscious awareness. *Trends in Cognitive Sciences*, 15, 365–373.
- Lurz, R. W. (2001). Begging the question: A reply to Lycan. *Analysis*, 61, 313–318.

- Lycan, W. G. (2001). A simple argument for a higher-order representation theory of consciousness. *Analysis*, 61, 3–4.
- Lycan, W. G. (1996). *Consciousness and experience*. Cambridge, MA: MIT Press.
- Neander, K. (1998). The division of phenomenal labor: A problem for representational theories of consciousness. *Notis*, 32, 411–434.
- Rosenthal, D. (1997). A theory of consciousness. In N. Block, O. J. Flanagan, & G. Guzeldere (Eds.), *The nature of consciousness* (pp. 729–755). Cambridge, MA: MIT Press.
- Rosenthal, D. (2002). How many kinds of consciousness? *Consciousness and Cognition*, 11, 653–665.
- Rosenthal, D. (2005). *Consciousness and mind*. Oxford: Oxford University Press.
- Schwitzgebel, E. (2008). The unreliability of naive introspection. *Philosophical Review*, 117, 245–273.
- Textor, M. (2013a). Brentano on the dual relation of the mental. *Phenomenology and the Cognitive Sciences*, 12, 465–483.
- Textor, M. (2013b). Brentano (and some neo-Brentanians) on inner consciousness. *Dialectica*, 60, 411–432.
- Tye, M. (2000). *Consciousness, color, and content*. Cambridge, MA: MIT Press.
- Vision, G. (1998). Blindsight and philosophy. *Philosophical Psychology*, 11, 137–159.
- Weiskrantz, L. (1997). *Consciousness lost and found*. Oxford: Oxford University Press.