

# Consciousness, Intentionality, and Self-Knowledge

## Replies to Ludwig and Thomasson

Charles Siewert  
Department of Philosophy  
University of Miami  
PO Box 248054  
Coral Gables, FL 33124  
U.S.A.

csiewert@miami.edu

Copyright (c) Charles Siewert 2001

PSYCHE, 8(09), October 2002  
<http://psyche.cs.monash.edu.au/v8/psyche-8-09-siewert.html>

KEYWORDS: phenomenal consciousness, mental reflexivity, higher order thought, rational agency, intentional feature, first-person knowledge, introspection.

REPLIES TO: [Ludwig, K. \(2002\). Phenomenal Consciousness and Intentionality, \*Psyche\*, 8\(08\).](#) and [Thomasson, A. \(2002\). Two Puzzles for a New Theory of Consciousness. \*Psyche\*, 8\(03\).](#)

ABSTRACT: Both Ludwig and Thomasson question my claim that many phenomenal features are intentional features. Further, Ludwig raises numerous objections to my claim that higher order mental representation is not essential to phenomenal consciousness. While Thomasson does not share those objections, she wonders how my view permits me to make first-person knowledge of mind depend on phenomenal consciousness. I respond to these challenges, drawing together questions about the forms of mental representation, the phenomenal character of sensory experience, rational agency, and introspection.

## 1. Introduction

Kirk Ludwig and Amie Thomasson address related issues concerning my views about phenomenal character and intentionality in their thoughtful comments on *The Significance of Consciousness* (Siewert 1998, all page citations will be to this book). So I would like, in the interest of economy, to combine my responses to them. While my

differences with them arise against the background of considerable agreement on fundamentals, such disagreements as we have raise questions that are far from minor.

On a general level, Ludwig's and Thomasson's criticisms challenge me to make it clearer how I am to harmonize two seemingly discordant tendencies in my views. On the one hand, I insist that phenomenal consciousness must not be consigned to the status of "raw feel." I object to accounts that separate mind into a phenomenal aspect and a representational/intentional aspect, and argue that the phenomenal character of experience is intimately bound up with understanding and the exercise of intelligence. Meanwhile, not only do I oppose attempts to swallow up consciousness in theories of intentionality, and, as part of this, reject the notion that some sort of meta-representation is (non-trivially) sufficient for consciousness-I do not allow that higher order mentality is even necessary for conscious experience. But the worry is that I may separate phenomenality from intentionality in some ways *too little* and in others *too much*. "Too little": because, in my zeal to rescue phenomenality from being reduced to brute sensation, I appear to run afoul of my own strictures against "consciousness neglect" (Thomasson). Or I ignore differences in the phenomenal character of experience that aren't differences in intentionality (Ludwig). "Too much": either because there is good reason to think phenomenality has higher order intentionality or reflexivity built into it (Ludwig), or because my segregation of phenomenality and reflexivity renders doubtful my belief that phenomenal consciousness is key to introspective knowledge (Thomasson).

I am still convinced I have articulated a coherent conception of the relation between phenomenal consciousness and intentionality, which provides suitable material for showing how consciousness enables us to know our own minds. But the challenge to remove apparent tensions in my views will be met only if I can respond adequately to questions Thomasson and Ludwig raise about two theses central to the last third or so of my book. The first of these theses I'll call "PI": "Phenomenal features are Intentional." This says that most, if not all, of the phenomenal features we (ordinary human beings) have, in thinking consciously and having visual experience, *are* intentional features. PI is the source of the concern that I separate phenomenality and intentionality too little. The second thesis I'll call "RI": "Reflexivity is Inessential" (for phenomenality). This says that, for having a phenomenally conscious experience, it is not necessary that one have a correlative mentally self-reflexive feature-i.e., an intentional feature whose conditions of truth or accuracy include one's having that experience. (To put it another way, it is not necessary that every phenomenally conscious experience involve a "higher order" representation of itself.) RI gives rise to concerns that I separate the phenomenal and the intentional too much.

I will start (in Section 2) by addressing the concerns raised about PI-mainly by clarifying the notion of intentionality I employ in it. In Sections 3-5 I'll turn to a defense of RI-here I will say a few words about the relation of conscious experience to concepts, memory, and rational agency. Finally, in Section 6, I will comment on how I propose to treat problems of "introspective" self-knowledge in a way consistent with my position.

## 2. Are Phenomenal Features Identical with Intentional Features?

Let me first sketch the defense of PI I give in the book. It draws (as does my case for RI) on a certain conception of what is sufficient for a feature to be an "intentional feature." A feature is an intentional feature if it is one in virtue of which its possessor is assessable for truth or accuracy. (pp. 188-192) The argument then, stripped to basics, is this. If that is indeed a sufficient condition for classifying a feature as intentional, the phenomenal visual features we can identify in thinking about the ways it seems to us for it to look as if things are shaped and positioned thus and so, are always (or almost always) intentional features. One need add nothing to their possession-no interpretation, no "meaning-giving act"-for their possessor to become assessable for accuracy with respect to them. At least, my search for plausible candidate "interpreting conditions" finds none that hold up. By a similar style of argument, I reach the conclusion that the phenomenal features we have in consciously thinking-whether in images or not-are features in virtue of which we are assessable for truth or accuracy, hence, intentional features. Now (as Thomasson notes) I claim no more than to have articulated a sufficient condition for a feature to be intentional (not a necessary-and-sufficient condition). Further-I leave open the possibility that there are or may be phenomenal features that are not intentional features, and I explicitly acknowledge that we have intentional features that are not phenomenal features.

Now one of the problems Thomasson raises is this: how can I maintain PI, and say some (many, maybe all) visual phenomenal features *are* intentional features? Does this not make me flunk my own test for realism about phenomenal consciousness, by denying the possibility of Belinda, whose blindsight gives her visual representation (hence intentionality), but without phenomenal visual experience?

I suspect that Thomasson's criticism of my identifying phenomenal with intentional features reflects a terminological problem. (I thank David Chalmers for his efforts to help me be clearer about this.) So I want to start by making a couple of points about the way I understand 'intentional feature.' Again, I say it is enough for a feature to be an intentional one, that it be a feature in virtue of which its possessor is assessable for accuracy or truth. And by this, I mean that it is enough that, when some x possesses this feature (as expressed in some proposition of the form '...x...F...') and certain conditions obtain, then something correlative of the form 'what x F's is true (or accurate),' or 'the way it F's to x is accurate (true)' follows necessarily. So: 'believes that Gore won the last presidential election' is an intentional feature, because there is a ("non-interpreting") condition such that, if it obtains, and x believes Gore won the election, it follows that what x believes is true (namely: that Gore won the election). And 'looks to x as if this circle is bigger than that one' is an intentional feature, since there is a (non-interpreting) condition (namely, that of this circle's being bigger than that one) such that, if it obtains and it looks to x...etc., then it follows: the way it looks to x is (in some respect, to some extent) accurate.

Let's call this (my) conception of intentional features, the "assessable-in-virtue-of" conception. And let's compare this with another one (perhaps more prevalent) that I'll call the "content conception." On this conception, an intentional feature is just the feature of having such and such an intentional (or representational) *content*. The idea is this. We may say a mental state has a certain content if it has certain conditions of truth or accuracy, or more generally "conditions of satisfaction"-or has these "under an aspect." And then we may say that a state's having an intentional feature is its having intentional content. (And we may say perhaps that a *person's* having an intentional feature is his or her having a state with a certain intentional content.) Notice that this "content conception" of intentional feature and my "assessable-in-virtue-of" conception are not just the same, and we might be inclined to individuate intentional features differently, depending on which conception we employ. On my conception, it will be natural to treat *believing that Bush will run for a second term* and *supposing that Bush will run for a second term* as distinct intentional features. But on the content conception, it will be natural to say that the state of believing that Bush will run...etc., and the state of supposing that Bush will run...etc., since they have the same content (that Bush will run for a second term), consist in the possession of the same intentional feature, though they differ in respect of the "mode" or "attitude type" they associate with that content/intentional feature. (For believing is different from supposing.) But, I think, which conception of intentional feature we employ is largely a terminological matter.

It should now be clear that when I say (at least some) phenomenal features are intentional features, I am not employing the "content conception." So I do not mean to say that possession of phenomenal features consists just in having a state of mind with a certain intentional or representational content-as if having a conscious visual experience with a certain phenomenal character were nothing more than having a certain representational content in mind. But it would also be a mistake to take me to say that having a conscious visual experience with a certain character is nothing more than having a certain content in a certain (say, visual) "mode." For one may understand "visual mode" so that a state can be in the visual mode even if it is not an occurrence of phenomenal consciousness. But I reject the view that a phenomenal visual experience is nothing more than a state with a certain type of content, had in some mode, of a kind that may pertain to *non*-phenomenal states. I do not think of phenomenal vision as what you get when you put a certain type of (possibly non-phenomenal) content together with a certain type of (possibly non-phenomenal) attitude or "psychological mode." I do think that if it looks to Sam as if one circle's bigger than another, he has a feature which is at once phenomenal and visual and intentional, but that is only because I take *looking* here to be an essentially phenomenal affair. (Blindsighters' stimuli can't *look* any way to them, in this sense.)

So much for my preparatory points of clarification. Now, how do these help me with Thomasson's criticism? I read this as taking the form of a dilemma that appears to arise for me once we take into account the distinction between the 'is' of identity and the 'is' of predication. Suppose I say that: (1) for some phenomenal feature F that a consciously sighted subject (Connie) has, but the corresponding blindsight subject (Belinda) is supposed to lack, and some intentional feature I, F is identical to I. Then, since I admit that Belinda has visual intentionality directed at what is in her blindfield, I am (allegedly)

committed to: (2) Belinda has the very phenomenal feature that, *ex hypothesi*, she lacks. And so (apparently) I contradict myself. Suppose I try to escape this unhappy result, and switch from the 'is' of identity to the 'is' of predication. So (3) I predicate, of the phenomenal feature F, that it is intentional. But then (4) I do not, despite my indication to the contrary, really challenge the traditional taxonomy of the mental into the phenomenal (or "qualitative") and the intentional (or representational) after all.

To start with: I do not really see the second horn of the dilemma as an alternative to the first. For it seems to me that if I *predicate* 'is an intentional feature' of Connie's phenomenal feature, then there is some intentional feature that this phenomenal feature is identical with—namely, itself. Unless perhaps we say that the phenomenal feature is only contingently an intentional feature. But I do not wish to say that. So it seems to me as if I am stuck with taking the first horn. But this does not worry me too much, since I don't think that (given my conception of intentional feature) (1) really does commit me to (2). That is, if I employ that "assessable-virtue-of" conception of intentional features rather than the "content conception" I can consistently say that what Connie has and Belinda lacks, when it seems to the former as it does for it to look to her as if there's greyish blurry patch on her left, is a feature that is both phenomenal and intentional. Though Belinda may have some sort of bona-fide visual intentionality with regard to her left-field, she and Connie do not have entirely the same intentional features on the "assessable-in-virtue-of" conception, even if they have just the same intentional features on some other conception—such as the content conception.

The key point is this. One will convict me of inconsistency by seeing me as committed to saying that Belinda has just the same intentional features as Connie, only if one assumes that I hold either: (a) having states with the same intentional or representational content is sufficient for possessing the same intentional features; or (b) that having states with the same (not necessarily phenomenal) content in the same (not necessarily phenomenal) mode is sufficient for having the same intentional features. But now we can see that I am committed to neither (a) nor (b). I can say that Connie differs from Belinda in respect of a feature that is at once phenomenal and intentional, because I am neither employing the content conception of intentional features, nor am I assuming that all intentional features can be analyzed into a "content" portion and a "mode" portion, neither of which is essentially phenomenal.

Notice also, that to maintain my position I do not need to argue for the idea scouted by Ludwig, that sameness or difference in contents of our perceptual states is always sameness or difference in the states' phenomenal character. Since I can allow that intentional features may differ in ("modal") ways that cannot be reduced to differences in content, I can have Belinda and Connie differ in phenomenal-intentional features without trying to trace this to a difference in the content of their perceptual states. I do acknowledge the interest of the issues Ludwig raises in this connection, concerning the limits to which different sense modalities may share content. I just do not see myself as needing to take them on, in order to maintain that some (many or all) of our phenomenal features are intentional features.

Maybe I have abetted confusion here by proclaiming that phenomenal features are intentional features without dealing with issues that arise once we put some of the notions of intentional content, object, and mode into play that have figured so much in discussions of intentionality. I held these notions and associated controversies at arm's length because I suspect they are often more trouble than they are worth, and because I thought I could leave them unresolved while making the central point I wished to make about consciousness in Chapters 7 and 8—that many phenomenal features are sufficient to give their possessors visual-spatial, recognitional and conceptual capacities. And indeed it seems the worries Thomasson and Ludwig make for me do not threaten (and are not intended to threaten) my central arguments that possession of visual phenomenality is sufficient to make one assessable for accuracy with respect to vision and that possession of noetic phenomenality is sufficient to make one assessable with respect to thought.

Perhaps my discussion would have been clearer if I had articulated my position while avoiding the jargon of intentionality altogether. Or, if this seemed unwise, I should have made more complete and explicit a stance that situates me with respect to the complex controversies now couched in terms of 'content' and 'representation.' I will say that, if I were forced to employ the content conception of intentional features, I think I could hold onto the essentials of my position, saying (as Ludwig and Thomasson suggest), that the possession of certain phenomenal features entails (but does not consist in) the possession of certain intentional features.

Now I will venture some comment on one more issue Ludwig raises regarding the individuation of phenomenal and intentional features. Why, he wonders, should we not individuate phenomenal features by reference to particulars, so that the way it seems to *me* to have the experiences I do could never be just the same as the way it seems to someone else to have the experiences he does? First, I think that, to the extent I am considering phenomenal *features*, I am conceiving of "repeatables"—what is of a nature to be attributed on more than one occasion. And further, at least some phenomenal features I can conceive of being had by more than one person. Moreover, when I set to conceiving of someone else having some of my phenomenal features, I find no obstacle in principle to the idea of a total phenomenal twin. At no point do I find a phenomenal feature I have such that I can say: "Now *that* one (a special "me-feeling"?) no one but I could possibly have." So I don't know how to identify the ineradicable phenomenal differences between individuals that Ludwig suggests we contemplate. And that is why I do not distinguish phenomenal features by reference to the particular person who has them. Maybe I am missing something in Ludwig's question, but that at least is where I would start my reply.

### **3. Ways of Seeming, Perspective, and Causal Content: Indications of Reflexivity?**

I now want to move on to discuss criticisms of my argument for RI (Reflexivity is Inessential), and the relevance of my position here for an account of self-knowledge.

The rationale for RI I present in the book goes roughly as follows. The question is whether every phenomenally conscious experience must be the target of some kind of reflexive intentionality. My starting point is to secure the recognition that, while we sometimes think about our conscious experience as it occurs, we do not do this continually. Here's where what Ludwig calls my "argument from examples" comes in. We know that sometimes we *attend to* our own silent speech, and our own visual experience, or we *report on* our own experiences, in ways that we find natural to say indicate we were then *thinking about* them or *thinking that* we had them. And we know also that often we do not attend to or report on them in this way. I think this shows that we do not always *consciously* think about our conscious experience as it unfolds. And I argue further, we have inadequate reason to posit the occurrence of some kind of *non-conscious* thinking directed at all the conscious experience that happens without our consciously reflecting on it at the time. I conclude that we should reject the idea that every experience is the object of some occurrence of higher-order thinking. So at least *that* kind of monitoring reflexivity is not essential to phenomenality.

However (as I acknowledge on page 202), this leaves standing the suggestion that some associated higher-order *dispositional belief* is essential to phenomenality. And, as both Ludwig and I point out: wherever one has recourse to first-person knowledge for examples of experiences not consciously reflected-on at the time they happen, one will have at least some reason to suppose that one did form a contemporaneous dispositional belief about the experience in question. For that seems to be a reasonable (though not, I think, indisputable) inference from the fact that one is now (retrospectively) thinking about the experience.

So, what Ludwig sees as my "argument from example" is intended only as part of the argument for RI. Next come the considerations that figure in what Ludwig calls my argument from "conceptual deficit." Observe that the higher-order beliefs about our experience we articulate in words involve the use of concepts (concepts of e.g., seeing, hearing, looking, sounding, and feeling), the possession of which requires something more than the mere capacity to see and hear and feel. What more is required? Plausibly: certain kinds of inferential capacities, and a commitment to relevant truisms (pp.202-3). And one could reasonably believe that one-year-old human children (and many non-human animals), lack these capacities and commitments, and hence, lack the concepts in question; thus they don't have the sorts of higher-order beliefs we (who possess these concepts) do have. Now, whether or not we *know* just how to describe the cognitive life of small children, dogs, monkeys, and so on, in this or that respect, we can know this. Our reasonable hypothesis that they do see, even while they lack the conceptual resources for higher order beliefs of the sort we possess, does not commit us to the proposal that they are all high-grade spontaneous blindsighters, responding with flexibility and finesse to visual stimuli that look no way at all to them. Therefore, ordinary higher order beliefs of the sort we may reasonably think young children lack should not be considered essential to phenomenal visual and aural experience.

But my argument does not rest here. For this leaves the suggestion that maybe toddlers (and monkeys, dogs, etc.) have pre-linguistic concepts of their experiences--distinct from

those we have and express in natural language--and these *primitive* concepts figure in higher-order beliefs they have about their experience, which are essential to phenomenal consciousness. Or maybe phenomenality requires them to enjoy a mental reflexivity not of a "conceptual" sort at all, properly speaking, but "non-conceptual" in character (in something like the way sensory representations of particular places are non-conceptual). I do not claim to show that babies have no "baby concepts" and "baby beliefs" about their own experience (or that dogs have no such "doggy" mentalistic notions). My point is that, we will have warrant for denying RI, and saying phenomenality always comes with reflexivity, only if we make it sufficiently clear either: what the hypothesized *primitive concepts* of mental states are, and why we should suppose babies and dogs must have them if they have phenomenal experience; or, alternatively, what *non-conceptual* higher-order representations are, and why we should suppose that all who have conscious vision, hearing, etc, must have *these*. I contend that neither condition has been met. Moreover, the prospects for meeting them seem to me dim enough that I conclude we have more warrant for denying that mental reflexivity necessarily attaches to all phenomenally conscious experience than for affirming that it does.

Once we take into account the structure of my argument for RI as just summarized, it seems that the real focus of disagreement with Ludwig is this. Ludwig thinks that there are several good reasons to maintain (contrary to RI) that reflexivity is essential to phenomenality, even without clarifying what primitive higher-order concepts are, or what non-conceptual higher order representations are. I disagree with him on this, but I admit that in my book I do not address the specific reasons he offers for holding that reflexivity is essential to phenomenality. I will now try to do so.

Ludwig first finds some indication that reflexivity is built into phenomenality in my "seems" talk--the locutions I employ to identify phenomenal features as "ways of seeming." If, when I wish to speak of the phenomenal feature I have in feeling pain, I say it seems to me as if I feel pain, does this not commit me to linking phenomenality and reflexivity? Its seeming to me as if I feel pain is some sort meta-representation of myself as feeling pain, is it not?

A problem here is that I do not myself report the possession of such phenomenal features by saying, "it seems to me as if I feel pain." The locution I favor is: "It seems to me as it does (or the way it does) to feel pain." I do not deny that this is equivalent to "It seems to me the way it does *for me* to feel pain," but I don't think it is clearly equivalent to "It seems to me as if I feel pain," if that latter is taken to attribute some mentally self-reflexive feature. I introduce my "ways of seeming" talk (pp. 86ff) to report first-person knowable differences, such as can obtain only among conscious experiences one has (differences in their phenomenal character), along with the stipulation that this is not to be taken to entail that some sort of reflexivity is packed into the phenomenal character of experience (p. 91). If readers find it irresistible to interpret my "seems" talk in manner that begs this question, I should try to find a different linguistic device to do the work.

But maybe we should see reflexivity in all sensory phenomenality for the reason Ludwig goes on to suggest--that the latter is closely bound up with the perspectival nature of



perception. Here I do acknowledge that the phenomenal character of visual experience of objects varies with the subject's location relative to objects seen. But I do not think this requires me to say that the phenomenal character of the experience somehow represents the subject as having that experience. Yes: it seems to me now the way it does to see these things *from here*, and that is different from the way it would seem to me to see the same things from over *there*. However: this does not entail that the way it now seems to me *represents* me as visually experiencing things from here-it does not entail that (as Ludwig puts it) the visual experience "represents itself as a state of the person whose state it is."

Perhaps however, RI is disproven by attention to the way causal conditions figure essentially in perception. Ludwig points out that there is more to seeing than having a visual experience, and what is additionally required for seeing is that one's experience be caused in the right way. And, he argues, what's required for seeing is required for the *veridicality* of experience, and so has to do with its representational content. Finally, the experience not only thus represents itself as having been caused in the requisite way; this self-reflexive aspect of its content is also in some way present in its phenomenal character-the way it seems to have it. Hence the phenomenality of perceptual experience is inseparable from a kind of reflexivity. Now what do I have to say to this?

Consider someone to whom it *looks* as if p is the case. We should distinguish between:

- a. The additional conditions that must be satisfied if the person is to *see* that p is the case.
- b. The additional conditions that must be satisfied if it is to look *accurately* to this person as if p is the case.

Now I want to concede that the (a) conditions-the conditions of perceiving -require that one's visual experience (its looking to one as if p) be *caused* in the right way. And the way the term "veridicality" is usually employed by philosophers, it does seem correct to say that an experience is "veridical" only if these conditions are met. But (b)-the conditions of "accuracy," as I understand them-don't include this causal element. That is, even if your experience is the product of a "deviant causal chain," it can, in principle, accurately represent the way your environment is. At least, we can understand such a use of the term 'accuracy.' Now, if we talk about the "content of experience," it is not clear to me that anything should be "read into" the content of the experience in this sense, other than what has to do with the accuracy conditions of the experience. Granted, the right causes have to be in place, if one is to see something is so, and not merely have an accurate visual experience of something's being so. But that doesn't show that these causal conditions are somehow represented in the experience. And I do not think that the phenomenal character of visual experience should be described by saying that the experience presents itself as if it is caused by the scene it represents. The fact that causal conditions are essential to perception does not give us reason to believe that there is some kind of reflexivity inherent in the content of a perceptual experience, or its phenomenal character.

Ludwig also proposes that we can recognize the inseparability of phenomenality from reflexivity once we reflect on the fact that, for any phenomenally conscious experience, there is *something that it is like for one* to have that experience. How, he wonders, can there be something it's like for our (allegedly unreflective) one-year-olds to have visual experiences of stimuli (as contrasted with mere blindsight) unless we suppose that the little consciously sighted tykes are aware of their "conscious visual fields"? And if they have such awareness, this would seem to constitute some kind of mentally self-reflexive state.

I would agree that children have phenomenal vision only if there is "something it is like" for them to see. And this, only if they are (in some sense) aware of the "field" of visual stimuli. (Not their "conscious visual fields" though-visual *fields* are not conscious, visual experiences are.) And I think a first-order awareness of visual stimuli-i.e., these look some way to the children-is sufficient for such an awareness of the field. Generally, I am not convinced that what has been called the subjectivity of experience, or (not much less obscurely) what it is like for someone to have experience, must be interpreted in terms of some second-order awareness of the experience. Notice that, in order to resist this interpretation, one need not claim to be able to imagine what it's like to be an experiencer who lacks a second-order awareness of experience. (Maybe that is self-defeating.) Instead, one can recognize that, in maintaining children to be without the concepts they need to form certain higher order beliefs about their visual experience, one is not committed to thinking they lack what we supposed the spontaneous blindsighter to lack, when we supposed there was no way her visual stimuli looked to her. This recognition does not require the (problematic) imaginative projection into young children's experiential life that may seem called for, when we try to imagine what it is like for them to experience things as they do.

## 4. The Argument From Memory

Ludwig also claims to find a necessary link between reflexivity and phenomenality by arguing roughly as follows.

- a. If you have a conscious experience at a certain time, then you are able to remember this at some subsequent time.
- b. If you can remember your conscious state, then you know you had it when you did.
- c. If you know you had it when you did, you represented yourself as having it at the time you had it.

More roughly still: (a) conscious entails memorable; (b) memorable entails known; and (c) known entails represented. Thus a conceptual connection between consciousness and memorability reveals the inseparability of consciousness and higher order mental representation.

I do not find this sufficient to establish a necessary connection between reflexivity and phenomenality. I have some doubt about each of (a), (b), and (c), but I will focus my criticisms on the first step, the "conscious --> memorable" step.

First: I do not think it is somehow primitively obvious that (a) is conceptually true. To decide whether we want to accept (a), we need to explore a little just what sort of relationship there seems to be between consciousness and memory. Do our intuitions really support the claim of a conceptual connection between consciousness and specifically *higher order* memory-memory that involves representational or intentional states whose conditions of truth or accuracy include the rememberer's possession of certain states of mind? We need, in considering these matters, to distinguish clearly the following two claims:

1. At time  $t$ ,  $x$  has conscious experience of object  $O$  --> At some time subsequent to  $t$ , (provided  $x$  still exists)  $x$  is able to remember  $O$ .
2. At  $t$ ,  $x$  has conscious experience of  $O$  --> At some time subsequent to  $t$ , (provided  $x$  still exists)  $x$  is able to remember that  $x$  had an experience of  $O$ .

My first point is this. We may be drawn to the idea that conscious experience and memory are inseparable, at least in part because we find it hard to deny (1) for something like the following reasons. It is hard to suppose that one might have a conscious experience of some object at a given time, without, at any subsequent moment at which one exists, even briefly being able to remember the object then experienced. This seems to me to have something to do with the need for a kind of continuity in an experience of an object. The thought is this: if my experience of an object at  $t$  did not involve even a fleeting capacity to remember that object at time  $t + n$  (for some  $n$ ), it is hard to see how I could experience the same object over time, that is, how I could experience it *as* the same-how there could be object constancy in experience.

But notice, so far no mental reflexivity is involved. Thus our attraction to (1) does not support (2). And if our acceptance of (a) is to get us the consciousness-reflexivity connection, we need to see it as entailing (2), which it is natural to interpret as linking consciousness to a higher order memory, and thus to mental reflexivity. But now, what support is there for (2)? I cannot say that its truth strikes me as obvious on reflection.

In conversation, Ludwig has suggested the following way of eliciting a conviction that might seem to favor (2). Consider again one-year-olds. Still, it seems that, even if their experience is wholly inaccurate or hallucinatory on some occasion, they could rightly be said to remember, or have some memory of, something from that time. But if, *ex hypothesi*, there is no object-beyond-the-experience for them to remember (they were hallucinating), what is left for them to remember, but the experience?

This might be taken to show (against me) first, that assuming small children are capable of having hallucinatory episodes and remembering something from them, we need to make room for some kind of meta-representation in their minds. Second, if we combine this consideration with reflections about object-constancy in experience we might seem

to get the result that any subject capable of experience "as of objects" also has higher order memory of those experiences.

In response I want first to note this. If the experience of object constancy requires some kind of memory of the object experienced, we could plausibly say that, in the case of *hallucinatory* object constancy, we have neither a capacity to remember the object experienced (there is none), nor a capacity to remember that one had a (hallucinatory) experience. Rather what we have is a capacity for an experience that is *like* remembering-only non-veridical (hence not really remembering)-a kind of "quasi-memory."

Alternatively we might describe the situation along these lines. We can recognize that someone might have a non-veridical sensory experience of an A, and then form, from that experience, a mental image or other representation of an A. Now what do we call the faculty by which we form an image of an A "from" or "on the basis of" an experience had of an A? It is a kind of memory, I suppose--it involves somehow "holding" or "preserving" the original experience "in memory". But does that kind of preservation necessarily require forming a higher order representation of the original experience, some representation that is true or accurate of that experience? I don't think it does. We could still describe the subject's current image of an A as one formed "from memory" of the earlier experience of an A--maybe even saying that image *is* a memory of the earlier experience, an "experiential memory." But then we are not interpreting 'this image is a memory of an earlier experience' in such a way as to entail 'this image is a higher order representation of the earlier experience.'

The image experience may be, in a sense, a *re*-presentation of the earlier experience from which it is formed--that is, it is an experience *like* the first, had *because* of the persistence of the (effects of the) first. And so we may say the second experience "copies" the first. But it does not, for all that, necessarily involve mentally self-reflexive intentionality. What makes the later mental image accurate or inaccurate (if anything) is whatever makes the original experience accurate or inaccurate. Note: one representation may be a copy of another (and a better or worse copy of it), without the copy being a *representation* of the original. I take it a copy may be a *copy* of the original--but not a representation of it. (To borrow one of Ludwig's examples: a globe may be a model, and a representation, of the earth. And that globe may be a model for another. But this second globe is also a representation of the earth, and a copy--not a representation--of the first globe.)

Thus we can describe as "memory" that power by which experience (even hallucinatory experience) of an A is somehow "retained" so as to give rise to later representations of an A, without assuming that any higher-than-first-order mental representations are formed. Compare: presumably, we can say that dogs' memories enable them to learn from their experience. We may even say Bowser bolts from the room when Uncle Ed enters because Bowser *remembers how it felt* to get kicked by him. But the sense in which such learning involves the "retention" of experience in memory does not require the dog to form a higher order representation of its pain experience.

So: such necessary link between experience and memory as I can discover can leave the objects of experience and memory both on "ground-level"-without meta-representation cropping up. Even if cases of non-veridical experience incline us to say that the subjects' memories are "of" past experiences, not extra-experiential objects, what underlies this talk need be nothing more than the kind of memory involved in forming mental images *from* (and in that way "of") a similar past experience, or the memory that is involved in associative learning. And for that, there need be no "higher order" mental representation going on.

## 5. The Argument From Agency

Though Ludwig's argument from memory does not persuade me of a conceptual connection between phenomenality and reflexivity, it does raise and help focus a number of subtle issues about the relationship between experience and memory-issues that need addressing in an account of consciousness, but on which I confess I have only made a start. Similarly while I'm also not convinced by Ludwig's argument from agency, I find it raises an important challenge from which I hope to learn. The argument, I take it, goes more or less like this.

Reasonable candidates for conscious experiencers without mental reflexivity (such as one-year-old children) are also justifiably regarded as rational agents. If they are rational agents, then they possess conditional beliefs with action-types in their scope (e.g., Sally believes that reaching for the bobble will bring it about that...). If children have such instrumental beliefs about (and hence concepts of) their actions, then they have concepts of (and hence can form beliefs about) beliefs and intentions. (For the concept of an action is the concept of movement rationalizable by belief and intention.) And if they have such concepts, it seems reasonable to attribute to them concepts of their own sensory experiences also. Or, if we withdraw attributions of both types of concepts, and treat children as non-agents, we should perhaps begin to doubt they have phenomenal consciousness as well.

Maybe, as Ludwig says, this argument gives us some reason not to deny concepts of seeing, hearing, sounding and looking to one-year-olds. But it does not really undercut my argument for RI. I can grant that the reasons I offer for thinking one-year-olds lack the concepts of seeing, etc. that figure in natural language expressions of our higher order beliefs are inconclusive. And I could grant that Ludwig finds serious reasons for the contrary view in his remarks about agency. Still, the thesis that children lack the inferential capacities and associated beliefs requisite for forming higher order beliefs about their visual and aural experiences is not an unreasonable one, even if ultimately the balance of reasons tells against it. And we may hold this thesis, while stoutly maintaining that children see things, but *without* committing ourselves to saying their vision is only a kind of blindsight. So the crucial moves in my argument for RI still stand.

Nevertheless, I think the points Ludwig raises about agency are interesting, and I need to come to terms with them sooner or later. Let me say a few words then about how I am inclined to view his argument linking agency and reflexivity.

First, I think we should seriously consider the idea that animals' possession of perceptual appearance, memory and desire can move them in some of the ways these would move genuine, full-blown rational agents/practical reasoners, though in the absence of the kinds of instrumental beliefs only the latter really have. Less-than-rational "proto-agents," behave, in some respects, as if they have those instrumental beliefs, but, precisely because they don't satisfy criteria for having the concept of belief or intention, we should hold that they don't truly have the beliefs in question. However, they do not generally have merely "as if" or "interpreter relative" intentionality, and they still count as "agents" of a sort-even if not "rational" agents. For, unlike thermostats and other artefacts to which we may (with varying degrees of utility) take the "intentional stance," they really have sensory experience, desires, and memories, and act on the basis of these.

Another approach we should consider is this. Let us attribute to primitive or proto-agents genuine instrumental beliefs about what kinds of acts will lead to what kind of outcomes (e.g., Sally wants the bobble, and believes that reaching like *this* will get her the bobble)-beliefs which "rationalize" and explain the acts in question. But still we might deny that their possession of such beliefs must involve their possession of the concepts of intention and belief. The agent, we may say, perceives her acts *as* acts of relevant types, which types are not to be verbally identified by us as mere movements, but (ineliminably) only relative to goals for which they are done (e.g., "reaching" is a motion (typically) for the sake of grasping). Nonetheless, to perceive her movements in this way the agent need not herself possess the concept that would be expressed by the phrases by which *we* identify the act-type. (For instance, she needn't conceive of reaching as a movement for the sake of grasping.) What then, enables her to think of reaching, so as to form instrumental beliefs? Maybe just this: her capacity to perceptually recognize her performance of the act-type in question, and adjust her execution of the act in ways that make her more likely to secure her goal. That she possesses such primitive, recognitional concepts of her acts in no way entails she possesses concepts-sophisticated or otherwise-of belief and intention. So this does not remove or weaken doubts about her possession of higher order beliefs about her perceptual experiences.

This is only a sketch of course, but generally, it is, I think, reasonable to explore the idea that we can make sense of certain kinds of agency, by attributing to agents practical thoughts about types of acts, without assuming that agents themselves have concepts of these types, such as we must employ to identify these types linguistically.

The upshot of my discussion of Ludwig's argument from agency is this. At most his argument establishes that there is more reason to say one-year-olds have concepts of their own states of mind, and thus higher order beliefs about them, than I acknowledge in my book. But this isn't enough to show that some kind of meta-representation is essential to the phenomenal character of sensory experience. Also, I don't think we should accept the premises of Ludwig's argument, without good reasons for rejecting the options for

understanding "primitive" or "proto" agency to which I have alluded. According to these views, either some agents altogether lack genuine (not merely as-if) instrumental beliefs that characterize practical reasoning, or else such instrumental beliefs as they have about their own doings require no concepts of belief and intention.

## **6. Reflexivity and First-Person Knowledge of Experience**

So I find that neither Ludwig's argument from perspective, nor his argument from the causal theory of perception, nor his arguments from memory or agency, compel me to recant my claim. I still think that reflexivity is inessential to phenomenality. But now what about the problem for first-person knowledge, or introspection, which Thomasson suggests this stance exacerbates? She wonders: how can I make phenomenal consciousness central to an account of first-person knowledge and warrant, when I think there is no necessity that phenomenally conscious states be objects of higher-order representations?

First, I am doubtful of the suggestion that if I rejected RI, and held that phenomenality entails reflexivity, I would be sitting pretty with respect to accounting for first-person knowledge and warrant. Suppose it is true that, having a phenomenal feature, I must in some way represent myself as having it. Does this still leave room for me to say sincerely (whether outloud or silently) that I have phenomenal features I lack, and fail to affirm I have the phenomenal features I in fact have? It seems that it should somehow-unless we want to defend a very strong form of infallibility and omniscience with respect to first-person judgments about experience. But then there is still an issue about how a special warrant attaches to the explicit first-person judgments we make, and how they can constitute knowledge.

Anyway, since I think I am stuck with RI, I suppose I should be less concerned about what its denial would buy me, and more concerned with what accounts of introspection are available to me once I accept it. I should note that I am very interested in the Shoemaker style account that Thomasson mentions, in which it is argued that reliable first-person beliefs about mind supervene on the possession of mentalistic concepts plus ordinary human rationality, and thus, first-person beliefs, because they are reliable, constitute knowledge. I am currently at work on a paper about Shoemaker's views. But I am not satisfied with his approach, partly for the reasons Thomasson mentions. Partly also, I am not happy with his way of spelling out the idea that rationality is inseparable from self-knowledge. For he works this out by appeal to claims of conceptual connections between first-order states of mind and higher-order beliefs about them, which I reject partly on account of the sort of considerations I appeal to in defending RI. However, I do agree that rationality presupposes self-knowledge. But I see a couple of problems with trying to use this rationality/self-knowledge connection to account for first-person knowledge of experience.

First, I am concerned that it either involves a hidden vicious circularity, or tacitly appeals to another (unexplained) first/third person asymmetry. For it seems to me that, if the fact of others' rationality is to provide warrant to my (third-person) beliefs about them, I need to have warrant for believing that they are rational in the relevant respects. But this seems to derive from my warrant for beliefs about their attitudes and experiences. Is the same thing true in the first-person case? If so, it looks like there will be something viciously circular about trying to account for one's possession of first-person warrant for beliefs about experience by appeal to the fact of one's rationality. For then, to warrant my first-person beliefs about attitudes and experience, I will need warrant for believing in my own rationality, which in turn will derive partly from belief in the correctness of my first person beliefs about attitudes and experience. (I thank Ludwig for prompting me, in another forum, to think about this concern.) We may respond to this by saying the first-person and third-person cases are different in this respect: the fact that others are rational can provide warrant to one's third-person beliefs about their attitudes and experiences, only if one has warrant for believing that they are rational. But the fact that one is oneself rational can provide warrant for one's first-person beliefs about mind *without* one's having warrant for believing in one's own rationality. But I find it doubtful that *this* first/third person asymmetry obtains, and one will wonder in any case why we should believe that it does.

Second, I doubt the Shoemaker approach can account for what I take to be a revealing fact about self-knowledge. Sometimes one's observers can have fairly good warrant for a certain belief about one's experience, where one has a first-person belief inconsistent with it, for which one has more warrant. And additionally, one has, in fact, no warrant at all for believing what others have most reason to believe about oneself in this instance. This kind of case arises where one knowingly deceives others about one's experience. Somehow first-person warrant is strong enough that it can enable this to happen, though it is not so strong that it precludes making errors that others can discover. I think it is a central challenge for accounts of self-knowledge to explain how this is so. I doubt that the Shoemaker style of account will do this, because it needs to appeal to the idea that there are special reasons to think first-person beliefs are reliable, which don't apply to third-person beliefs. And it seems this will help explain how the sort of epistemic situation I've mentioned can arise, only if it can somehow be made out that the behavior warranting third-person judgment is generally a less reliable indicator of the types of experience for which that situation can occur than first-person belief. But it does not seem that this assumption is somehow entailed by the subject's basic rationality. Anyway, it is far from clear it is true.

What then is *my* approach to accounting for introspective knowledge? I can only outline my proposal here; more detail is provided in my article "Self-Knowledge and Phenomenal Unity" (*Nous*, Vol. 35, No. 4, December 2001). My idea is that phenomenal consciousness is basic to an account of the mind's self-knowledge, not because some kind of accurate self-representation is somehow part of the structure of phenomenal character, but because a person's conscious thoughts and experiences normally display a kind of unity that is also an aspect of their phenomenal character. This ("phenomenal") unity of conscious thought and experience allows one to direct one's thought to one's experience,



by *attending* to it in a way in which one cannot attend to another's—a way in which one also could not attend even to one's own experience in pathological or abnormal conditions of dissociation. By means of such attention one can understand one's use of "demonstrative" terms (like 'this' and 'that') to refer to one's own experience in thought (as in the thought: "This is a painful feeling"). Such thoughts are distinct from those expressed by claims in which one attributes the same experiences to oneself using the first-person pronoun (such as: "I feel pain").

Now I would note a disanalogy between the way in which one attends and refers to one's own experience, and the way in which one attends and refers to objects in one's environment. In the latter case, we are able to understand reference through a direction of attention that depends on forms of sensory appearance, through which the objects referred to are represented—not by use of general concepts—but as occupying particular regions of space. However, one employs no such non-conceptual form of representation of the experience referred to, in order to understand which experience one refers to in thought. In this sense, there is, strictly speaking, no such thing as "inner perception" to direct one's thought about one's own experience.

However, it is, ironically, not our possession of a special "inner" sense, but our very *lack* of it, that helps explain how we are able to know our own experience. For since we lack this, we understand reference to our own experience not by *its* means, but rather through a direction of attention based on the unity of consciousness. But that does not permit us to understand our use of demonstrative terms like 'this' in situations where there actually is no object there to refer to (as can occur in, for example, a case of hallucination). So, I argue first, if we understand what we're saying at all when we purport to refer to our own experience, there actually is something to which we refer. And second, we would not be able to understand our reference to it, unless some way in which we were disposed to classify that referent, using general concepts, were correct. Thus if also (as I argue) there is a warranted presumption in favor of one's conviction that one understands what one is saying, then there is a warranted presumption that we correctly conceptualize our own experience in some way. Since the same sort of reasoning will *not* show that we each have a similar right to make judgments about *other* people's experience, this argument justifies the claim that we have a distinctive sort of knowledge of our own experience, and explains what it consists in.

Now does this approach give me any way of dealing with the challenge Thomasson raises? She wonders how I can account for my knowledge of my own conscious experiences that are not consciously thought about or attended to at the time I have them. It is a serious challenge, and my account, as just sketched, does not deal with it directly. But I think it can be extended in this direction. I would propose that I know what I'm thinking as I am thinking, what I am experiencing as I experience it, even without contemporaneous reflection, insofar as I am able to attend to these experiences appropriately. Now in some cases I will only be able to do this immediately afterwards, with short term memory, so as to make (conscious) judgments about them that are phenomenally unified with them. (Phenomenal unity, as I understand it, can hold between experiences that are not contemporaneous.) But, I would argue (analogously to the way I

did before), one would not be able to understand the demonstrative expression of reflective thoughts that are phenomenally unified with the just-past experience or thought they are about, unless in so thinking one could somehow correctly apply concepts to the "lower order" experience or thought. There is a warranted presumption that one can correctly conceptualize these phenomenally unified experiences and thoughts in their wake. For we must be able to "keep track of" temporal stretches of such thought and experience accurately, if we are to think these demonstrative "reflective" thoughts at all. And this, I would argue, yields a warrant for correlative first-person past-tense judgments attributing thought and experience, where similar considerations yield none for such third-person judgments. So the suggestion is this: as I was thinking and experiencing, I acquired the capacity to make future retrospective first-person judgments about them with warrant. This then can make it correct to say that I knew what I was thinking or experiencing at the time, even if I didn't then reflect on it.

Thomasson also raises the concern that, even if my style of account can show how we have some of the first-person warrant we do, it still may not account for the extent of such warrant for judgments about experience. For maybe it won't explain with what right we make the more nuanced judgments characteristic of phenomenological description. I acknowledge I have not yet dealt with this issue. I would like to say that we may be more vulnerable to mistakes when we try to talk with a higher degree of specificity and refinement about our experience than is usual, and thus we may then have less warrant for such reports of experience than for the run-of-the-mill variety. But still: the presumption that we can make sense of our own experience yields first-person warrant here as well. It's just that here the greater precariousness of our conceptual grip means that this warrant will not be had without greater caution taken in considering and avoiding relevant sources of error.

Such, at any rate, are the strategies of argument I propose to explore. I believe-though I will not try to argue the case here-that this style of account can cope with relevant challenges that others-like Shoemaker's-cannot. But clearly, my account requires much more elaboration and defense. I provide this sketch only to indicate the potential for ways of dealing with problems of introspective knowledge without either appealing to an essential reflexivity in phenomenal consciousness, or a "reliability is a consequence of rationality" approach.

I hope my responses here have given the reader some idea of how I see issues regarding phenomenal consciousness, intentionality, mental reflexivity, rationality, and self-knowledge as significantly and inevitably linked. I am particularly grateful to Thomasson and Ludwig for the impetus their remarks have given me to try to attain greater clarity on these issues, and how they hang together. The difficult questions they raise are among those I most need and want to answer in further developing my views.

## References

Siewert, C. (1998). *The Significance of Consciousness*. Princeton: Princeton University Press.