

Does Opacity Undermine Privileged Access?

A Critical Notice of *The Opacity of Mind* by Peter Carruthers (OUP 2011)

Published in the *International Journal of Philosophical Studies*
Vol. 22, No. 4 (2014), pp. 617–629.

[This is the penultimate draft. The final version is at tandfonline.com]

Timothy Allen & Joshua May

Abstract: Carruthers argues that knowledge of our own propositional attitudes is achieved by the same mechanism used to attain knowledge of other people’s minds. This seems incompatible with “privileged access”—the idea that we have more *reliable* beliefs about our own mental states, regardless of the mechanism. At one point Carruthers seems to suggest he may be able to maintain privileged access, because we have additional sensory information in our own case. We raise a number of worries for this suggestion, concluding that Carruthers’s new theory cannot clearly preserve the superior reliability of our beliefs about our own attitudes.

1. Introduction

Peter Carruthers has recently developed an innovative theory of how we acquire beliefs about our own minds as well as others’. This theory, defended most recently and extensively in *The Opacity of Mind* (2011), challenges the common view that the mechanisms used to acquire such beliefs are importantly different. Drawing on empirical work in cognitive science, Carruthers argues that we don’t have direct access to an entire category of our own thoughts—roughly what many include under the label “propositional attitudes.” Instead, our access to these thoughts is “interpretive.” One determines what one believes, desires, hopes, and so forth much in the way one figures out when others are in these states of mind. However, our access to our own attitudes is different in that we rely heavily on our own perceptions and sensations as clues—hence his Interpretive Sensory-Access theory of self-knowledge. Although Carruthers doesn’t mention it, this is the same basic theory of self-knowledge first proposed by Ryle in *The Concept of Mind* (1949/2009, ch. 6; see Byrne 2012).

While Carruthers gives this provocative theory an impressive defence, we argue that it has an undesirable consequence. So, instead of addressing Carruthers’s overarching argument, we draw out a problem for the theory that can be addressed separately. We focus

on an issue that isn't adequately addressed by Carruthers: whether the ISA theory can avoid a more sceptical conclusion about self-knowledge by preserving "privileged access" to our own minds—a desideratum for many epistemologists working on self-knowledge, even if not for Carruthers. In particular, even if we grant that the mechanisms for self-knowledge and other-knowledge are the same, we may retain more reliable access to our own minds simply due to the greater amount of sensory information available.¹ Yet we argue that this is not so easy to maintain. Greater reliability cannot obviously be grounded in sensory states which are rather fleeting. While Carruthers may be able to avoid this problem with his "outward-looking" account of how we access sensory states, this manoeuvre is fraught with difficulties. Thus, an internal tension remains between maintaining an asymmetry in reliability for two types of knowledge while positing a shared mechanism.

2. The ISA Theory

Carruthers's Interpretive Sensory-Access (ISA) theory states that we access our own mental states differently depending on their type. We have direct introspective access to *perceptual or sensory states*, which includes states of perception (such as visual and imagistic states), emotion, proprioception, and interoception. But we have more indirect "interpretive" access to *propositional states* (or attitudes), a category which encompasses all non-sensory mental items, broadly categorised as judgements and decisions (2011, p. 1). In an earlier article, Carruthers explains that judgements are "events of belief formation", and decisions are "acts of willing, or the events that create novel activated intentions" (2010, p. 78). He also uses a broad definition of *introspection* as (1) a higher-order process, meaning that we acquire awareness or beliefs about our mental states, and is (2) epistemically different in kind from the access that we have to the mental states of others (2010, pp. 76-77).

Carruthers proposes a model of mental state attributions in order to explain the central claims of the ISA theory (2011, pp. 1-2; 2010 pp. 79-83). There are a number of perceptual systems in the mind that broadcast sensory information to a set of conceptual systems that includes the mindreading faculty, which produces higher-order judgements about first-person and third-person mental states. Carruthers's theory, put into the language of this model, is that the mindreading faculty receives input only from these perceptual systems. However, it's important to note that Carruthers believes perceptual information can itself be conceptual—e.g. "we don't just see a round green object, we see an apple" (p. 3; cf. pp. 73-5). The result is that the mindreading faculty can only self-attribute propositional states by *interpreting* this perceptual or sensory information, just as it attributes propositional states to other people.

Thus, one's mechanism for attributing propositional attitudes to oneself is not different in kind from the mechanism for attributing such attitudes to others. The difference, instead, is in degree: in one's own case there is a richer source of sensory information, such as inner speech and mental imagery. For example, suppose that Sally's running shoes are old and falling apart. In self-attributing the belief that she needs to buy a new pair, her mindreading faculty has at its disposal: the tattered appearance of the shoes she's wearing, the feeling of tightness and pain in her ankles, the feeling of embarrassment she experiences

¹ We assume greater reliability here is a matter of having a greater *proportion* of accurate judgments produced, not merely a greater *number*. Thanks to an anonymous referee for pressing for clarification.

when wearing them in public, and the visual image of a clean and comfortable new pair of shoes.

In contrast, the mind can self-attribute perceptual or sensory states directly, without undergoing a process of interpretation. This is because perceptual states constitute the input used by the mindreading faculty, which is received directly from the world by the sensory systems. Carruthers thus asserts that the mindreading system will find it “trivially easy to self-attribute those percepts,” as he puts it in an earlier paper (2010, p. 80). For example, when the mindreading faculty receives a visual representation of a dog chasing a ball, it will be trivial for it to output the perceptual state, “I am seeing a dog chasing a ball.” Such self-attributions are simply a process of *recognition* or “quasi-recognition,” as the mindreading system does not need to make an interpretation or an inference to output the self-attribution (2010, pp. 80-81; 2011, pp. 51, 80). Consequently, Carruthers says that this process qualifies as a form of introspection.²

Having described the basic method of mental state attributions in the mind, Carruthers summarizes the three central theses of the ISA theory (2011, pp. 1-2):

- (a) there is a single mental faculty underlying our attributions of propositional attitudes, whether to ourselves or others;
- (b) this faculty has only Sensory-Access to its domain;
- (c) its access to our attitudes . . . is interpretive rather than transparent [i.e. non-interpretive].³

Carruthers goes on to provide support for his ISA theory by drawing on a broad range of evidence from cognitive science, neuroscience, and psychology. His overarching argument is that the ISA theory is able to consistently integrate a great deal of empirical findings with current theories about mental architecture, and that it provides a simpler explanation than the alternative theories, which posit two separate mechanisms for self- and other-knowledge.

Self-knowledge is especially interesting because it seems to have distinctive properties. So epistemologists will no doubt wonder at this point: What exactly are the epistemic ramifications of this model? At first glance, the ISA theory seems to make a rather radical proposal: we have no special method of accessing our own propositional states, such as desires and beliefs, which seems to suggest that our access to these attitudes is about *as reliable* as our access to the attitudes of others. In fact, this is especially apparent when considering the phenomenon of confabulation in which people inaccurately attribute propositional attitudes to themselves. This is the “the central, key, prediction” of the ISA theory, according to Carruthers: “since people are often misled when they attribute thoughts to others, there should be frequent instances of confabulation [in self-attributions] resulting from misleading behavioral or other sensorily accessible cues” (p. 6). This suggests that Carruthers’s theory presents a sceptical challenge to the standard conception of self-knowledge as involving what we’ll label *privileged access*—i.e. greater reliability in our

² See Zimmerman (2008) for a defence of a somewhat similar view—i.e. one that treats knowledge of one’s own thoughts quite differently from knowledge of one’s sensations.

³ Unlike many others, Carruthers uses the term “transparency” to refer to mental states that are directly available to us through introspection, as opposed to interpretation (2011, pp. 1-2). To avoid confusion, we’ll generally avoid the term “transparency.”

attributions to ourselves than to others.⁴ It's widely held among epistemologists that we have such privileged access to our own mental states—so much so that it is often considered something of a desideratum (cf. Byrne 2011). Even Carruthers himself sometimes seems concerned to capture something like privileged access. So it is important to know whether such a theory can accommodate this apparently special access we have to our own attitudes.

There may be a way for the mindreading system's self-attributions to be more reliable than its attributions to others, even if they share the same mechanism. While Carruthers tends to imply that non-interpretive access entails the denial of privileged access, the two are dissociable in principle (cf. Byrne 2012). So we can ask: While the ISA theory can explain why it *seems* to us that we have privileged access (ch. 2), is this an accurate appearance? This is a “puzzle” that Carruthers more directly addresses in earlier work (2010, p. 105). It's unclear whether he has since changed his mind or become less concerned to preserve privileged access. At one point, he writes that such an issue about reliability is “moot” (2011, p. 70) and “currently unresolved” (ch. 3, n. 11).

Regardless of where he currently stands, Carruthers still mentions a way that might allow us to address this decidedly epistemic matter. He notes early on that when self-attributing propositional states, we can “make use of the evidence provided by our own inner speech, our visual and motor imagery, our own affective feelings, and so on” (2011, p. 3). We have seen that self-attributing perceptual states is “trivially easy”, because the mindreading system simply recognises the sensory information that the subject is receiving, and self-attributes a perceptual state accordingly. Consequently, the subject's higher-order knowledge about her perceptual states matches the sensory input. This seems to provide a solid basis for propositional states to be self-attributed with potentially high reliability. While the process is interpretive, the mind has a rich variety of sensory information at its disposal, which is taken both from the world (such as visual information) and from the sensory systems in the body (such as proprioceptive information). Take our running shoes example: given that the mindreading system can access visual states, interoceptive states, affective feelings, and visual imagery in order to interpret that Sally believes she needs to buy new shoes, it is not surprising that its attribution is highly reliable. So far, so good.

3. Reliability Problems

Yet there are several reasons to worry that the ISA theory makes for less reliable self-attributions. We'll raise two worries on this front: one concerning Carruthers's own manoeuvre here and another that appeals to an independent problem.

3.1. Informational Overload

Despite one's mindreading system having access to a wealth of one's own sensory states, Carruthers does briefly note that there is no necessary connection between the presence of more sensory data and reliability. In some cases, attributions are easier to make when there is a “smaller set of cues,” as the inferential mechanisms can “become overwhelmed by data” (2011, p. 24).

⁴ It's not clear that Carruthers always uses the term “privileged access” in this distinctively epistemic way. But we'll reserve it for this use only (cf. Byrne 2011; 2012).

This suggests that sometimes one can attribute mental states to someone else more reliably than they can attribute mental states to themselves. For example, suppose you're driving down a busy, unfamiliar street at night, you're running late for some event, and you're paying attention to many different things: the car's brake lights in front of you, pedestrians trying to cross between the stopped cars, the traffic lights up ahead, the speed limit, the time, the directions you're trying to follow, and the things your friend is saying from the passenger seat. But these are only the external sensory cues; you're also attending to your inner speech and mentally picturing the embarrassing scene of your late arrival. Your calm friend sitting next to you may be able to more accurately ascribe propositional attitudes to you, at least at that moment when you're distracted and being bombarded with sensory information.

Perhaps this won't seem so problematic. This kind of first-person inaccuracy may be rare, limited to cases where the mind is overwhelmed with sensory information to such a degree that it cannot operate as it does in normal circumstances. The ISA theory could then hold that in normal circumstances the mindreading system is capable of reliable self-attributions, and that first-person attributions are generally more accurate than third-person attributions. While Carruthers may urge that self-knowledge isn't always highly reliable, privileged access comes in degrees, so self-knowledge may remain more reliable than other-knowledge. Still, informational overload reveals that more information about oneself does not necessarily entail more self-knowledge.

3.2. Fleeting Sensory States

Even if the previous worry doesn't succeed on its own, notice that the less sceptical take on the ISA theory relies on the claim that our access to our sensory states is stable and reliable. Carruthers attempts to support this claim by reviewing a wide range of empirical research on our access to affect (ch. 5). However, his conclusion is rather mixed: we have and make use of direct, non-interpretive access to the *valence* of our affective states, but we may not make much use of our "theoretically available" access to *emotional* states, and we lack non-interpretive access to the *strength* of such states (see pp. 154-5). Moreover, Carruthers admits that "many emotional states can be quite fleeting" (p. 134). Of course, he makes this claim by way of defending the reliability of our access to certain affective states. However, we suggest that this raises a general worry about grounding self-knowledge of attitudes in such fleeting sensory states.

Even if we grant the mixed conclusion—that we have non-interpretive access to some aspects of some sensory states—this does not necessarily provide the epistemic upshot: a stable base on which the mindreading faculty can generate reliable self-attributions of propositional attitudes. To make this worry vivid, we'd like to add further reason to suspect that many of our sensory states are fleeting and our access to them is not necessarily so reliable. Consider Eric Schwitzgebel's (2008) sceptical challenge to the idea that we can reliably introspect our current phenomenal experience. He argues that while we are able to introspect current conscious experience, this introspection is quite unreliable. Schwitzgebel interestingly uses this challenge to argue that our beliefs about the world (outside of the mind) are more stable and trustworthy.

There is some empirical evidence for this hypothesis based on systematic reports from ordinary people attempting to describe their inner experience (Hurlburt & Schwitzgebel 2007 for discussion). But let's just focus on the more intuitive arguments based on

uncertainty when attempting to introspect different aspects of one's own phenomenology, such as emotional experience, visual perception, and thoughts. Take the latter first: thoughts are so central to our mental lives, yet there is great disagreement about whether thoughts have phenomenology, and if so, whether such phenomenology has particular properties. Do thoughts consist entirely of imagery; are they accompanied by feelings or emotions; or do they go beyond feelings and images? As Schwitzgebel points out, if we did have a reliable capacity for introspecting our own phenomenology, then we should at least be able to agree about this basic aspect of conscious experience (2008, p. 257-9).⁵

The problem arises for even the clearest examples of phenomenology, such as vivid pains and visual experiences of bright colours. For example, it is difficult to determine the exact location of a pain or to describe its qualities. The same goes for our various emotions, such as anger, joy, fear, or disgust (Schwitzgebel 2008, p. 249). The important category of inner speech is no better. It seems obvious that we have it sometimes, but it does often seem elusive once we attend to it (p. 260). Arguments based on reports of one's own phenomenology are slippery themselves, but one can presumably recognize the difficulties in one's own case. As Wittgenstein might say, try the following experiment: Have someone say to you "The weather is nice today" and then repeat this in "inner speech." Discerning their similarities and differences is difficult to say the least.⁶

It may seem quite controversial to suggest that we cannot reliably introspect phenomenal states. These states seem more likely to be directly accessible, and thus more reliable, since they are apparently presenting themselves in conscious experience at the time of introspection. However, the foregoing considerations suggest there is something flawed in the idea that we have better access to our own current conscious experience than we do to the outside world. As Schwitzgebel vividly puts it: "The teetering stacks of paper around me, I'm quite sure of. My *visual experience* as I look at those papers, my *emotional experience* as I contemplate the mess, my *cognitive phenomenology* as I drift in thought, staring at them—of these, I'm much less certain" (p. 267). Even if we cannot settle the issue here, there is at least some *prima facie* support for the idea that one's awareness of phenomenal and affective states is less stable than one's access to one's beliefs, desires, and similar attitudes.

In any event, these considerations raise concerns for the apparent compatibility between the ISA theory and privileged access. If our access to sensory states is to be privileged, there must be a rather straightforward and reliable transition between our sensory states and our knowledge of them. According to the ISA theory, when the mindreading system receives sensory states such as an affective feeling as input, it is able to output a self-attribution of the feeling with ease, simply by recognising it. But there is some reason to think that the process is far from trivially easy, at least for states that are essentially

⁵ At one point Carruthers suggests the ISA theory implies that there are almost no conscious propositional attitudes (ch. 12.2). If this turns out to be true, then this way of motivating the problem may not be effective against Carruthers. Still, we include it because it helps to raise the worry independently of Carruthers's further thesis about the phenomenology of thought.

⁶ It might seem that the ISA theory makes no claims about what will happen when we consciously attend to the process of self-interpretation. Yet the model Carruthers develops should precisely allow for this. As he says about other-interpretation, which of course is meant to involve the very same mechanism as self-interpretation, this can "operate in a reflective, slowed-down, and partly conscious mode," as when we ask ourselves: "What is he up to? Is he trying to open that door?" (p. 70).

qualitative or phenomenal. Ordinary self-reports about the phenomenal character of such affective feelings differ widely. If self-knowledge of affective experience really is just a matter of the mind recognizing the sensory feelings it receives from the body, then it is unclear why the basic features of our own phenomenology are so difficult to access.

Perhaps this is not just a problem for Carruthers's ISA theory. The concerns about the reliability of our access to sensory states seem to apply generally, regardless of the theory one holds about the method of this access. By this reasoning, it's not the case that the ISA theory fails to preserve privileged access to sensory states, but instead that these states simply cannot be accessed with greater reliability. This may be true, but it does not follow that the assumption of privileged access is false generally. These considerations don't raise doubts about the reliability of our access to states that appear to be stable and easily accessible. Even though we tend to experience great difficulty in determining whether our thoughts are always accompanied by images, for example, we don't seem to have this difficulty in firmly grasping our beliefs and working through their subtler components. So, even if we accept that privileged access doesn't extend to fleeting sensory states, privileged access does seem to be true of propositional states, and other similarly stable states. This leads to a more concerning implication of the ISA theory. It seems unable to explain this reliability that we experience when accessing our beliefs, desires, and similar states. Since the central claim of the ISA theory is that the mindreading system relies on sensory input to interpret our propositional states, it is unclear how such a variable input could provide the basis for stable and reliable access to our propositional states.

Of course, there might be various ways for Carruthers to respond to the idea that our sensory states are more unstable and fleeting than one might initially believe.⁷ But we contend minimally that this, along with worries about sensory overload, raises an important challenge for the ISA theorist who seeks to preserve privileged access to propositional states. Moreover, it connects with another important piece of Carruthers's theory that may save him from a more sceptical approach, to which we now turn.

4. An Outward-Directed Solution?

Carruthers maintains that we acquire knowledge of our own sensory states by attending outward to the world, rather than by looking inward (chs. 4-5). He writes that "self-knowledge [of sensory states] can be reliably acquired from knowledge of the world as it presents itself to the subject", as the input to the mindreading faculty is simply sensory representations of properties of the world (pp. 79-80).

This idea is often referred to as "transparency," but Carruthers reserves that term for non-interpretive accounts of self-knowledge of propositional attitudes generally. So we'll follow him in calling such an approach "outward-directed." Whatever label one uses, the idea is one that largely originated in Gareth Evans's (1982) famous example: In answering the question "Do you think there is going to be a third world war?," you will typically attend to

⁷ Jakob Hohwy (2011) argues that the phenomenological uncertainty and disputes to which Schwitzgebel points are best explained by phenomenal variability rather than the unreliability of introspection. Yet even Hohwy's view seems to concede that there is a substantial amount of variability and instability in our sensory states, which could still make privileged access more difficult for the ISA theorist to capture.

precisely the same outward phenomena as you would if answering the question “Will there be a third world war?” (1982, p. 225). While Evans thinks this account is true of beliefs, as a number of recent epistemologists do, Carruthers argues that it is only correct for knowledge of one’s sensory states (p. 79).

Privileged access may thus be grounded in the idea that, in our own cases, we have a wealth of sensory information that is accessible via an outward looking process. Even if sensory states are fleeting when we look inward, as Schwitzgebel contends, this is not how the mindreading faculty works. It can simply recognise sensory information in one’s own case by looking outward. Thus, Carruthers could argue that this provides a stable foundation from which the mindreading system makes especially reliable self-attributions of propositional attitudes.

4.1 The Evidential Objection

There are several known problems for outward-directed approaches generally, and we content that some are especially problematic when applied to sensory states in particular. But it depends on which version of an outward-directed approach Carruthers takes. One option remains close to the original form of Evans’s example, but generalizes to apply to more than just belief: we determine whether we’re in mental state M with content C by determining whether C is true. Yet this is subject to an objection pressed most forcefully by Matthew Boyle (2011), which is roughly that there is no evidential connection between states of the world and states of the mind. An outward-directed mindreading system would rely on bad inferences, leading to unreliable self-attributions.

To see the force of this worry, we need to recognize that outward-directed approaches involve an inference or transition: one infers one’s mental states from considerations about whether its content is true. (Or if the process isn’t attributable to the whole person: one’s mindreading system outputs a belief about what mental state one is in based on input from the world, as it presents itself to the subject, concerning the content of that state.) In the case of belief, the inference can be expressed as the transition from P to *I believe P*, a step that Alex Byrne (2011) calls the “doxastic schema.” Yet these are not evidentially related. As Boyle has recently put it, “the inference is mad” (2011, p. 230). When making an inference, we can ask what grounds we have for holding the inferred proposition to be true. In the case of the doxastic schema, what are the grounds for holding that one believes P? Even if P is true, this provides little evidence that one *believes* that P. The evidential connection remains lacking, even if there is normally a tight causal connection between the two. Of course, when one sincerely recognizes or asserts that P, one then seems warranted in judging that one believes that P, but this recognition involves a step that the doxastic schema lacks, which is required to remain purely outward directed.

As we have seen, Carruthers takes an outward-directed approach to self-knowledge of sensory states, not beliefs. But this approach is subject to the evidential objection as well, even if we just focus on belief-like states, such as perceptions. To illustrate using Carruthers’s terminology, suppose you’re looking at a green apple in your hand. An outward directed approach would have you inferring that you see a green apple from the fact that there is a green apple. However, there is no evidential connection between the information about the apple and the self-attribution, *I see a green apple*. A better inference would be the reverse: from “I see an apple” to “There is one.” Thus, the ISA theory on this construal posits a mechanism for self-attributing sensory states that is patently unreliable. And the same goes

for any sensory state—from visual experiences to inner speech. Whether applied to propositional attitudes or sensory states, the doxastic schema is plausibly prone to error.

Now Boyle argues that we should adopt a “reflective” approach to self-knowledge of beliefs instead, according to which introspection involves “shifting one’s attention from the world with which one is engaged to one’s engagement with it” (p. 228). Perhaps Carruthers could simply apply this model to sensory states. For example, if one is tacitly aware that there is a green apple precisely due to seeing it, it seems one is warranted in reflecting on that awareness to self-attribute the corresponding perceptual state. But, as with the case of belief, this would involve an inward gaze, precluding a purely outward-directed account. One would be self-attributing the experience in part by reflecting on that very experience. Yet, in that case, the theory would remain subject to the problem of fleeting sensory states.

4.2. Problem of Limited Application

Even if we ignore the evidential objection, there is another problem with this outward-directed approach. It seems to fail for self-knowledge of states that aren’t like beliefs. In Evans’s case, the idea is that we self-attribute the belief that P by answering the question of whether P is true. What seems plausible about this idea is that there is usually a strong (even if not evidential) connection between P and one’s belief that P, at least when one is being rational.

But this connection is absent in cases other than those involving belief-like states, such as desires. It would be foolish to attempt to determine whether one desires to have coffee by checking to see whether one does in fact have coffee. Typically it’s the falsity of the content of my desire (assuming it’s propositional) that instigates having the desire in the first place. In response, one might modify the account, saying it’s outward-directed in a different sense: we determine whether we are in mental state M by determining whether we *should be in M* (cf. Way 2007). This has the potential to apply to mental states of various types, even ones whose function is not to accurately represent their contents.

However, even this kind of outward-directed account can’t apply to every kind of state (cf. Owens 2003). Common examples of temptation and weakness of will show that we don’t always want something to be the case even when we think we should. For example, someone suffering from severe depression may believe she should want to attend her friend’s wedding but lack any such desire. While one may be properly counted as being irrational or suffering from a mental malady, these cases certainly seem to occur.⁸

The general problem here seems to be that we don’t always have the mental states we think we should, especially when it comes to affect and motivational states. Yet some sensory states fit this bill precisely, including emotions and even inner speech. For example, Jonny may believe he shouldn’t recite suicidal or homicidal thoughts, but they occur anyway. The mindreading system cannot then determine that one is in such states by simply looking outward and addressing questions such as: Should one be in such states? We can easily cook up examples for emotions, such as embarrassment, to connect with our running shoes example. But the problem applies even to certain perceptual states, assuming it even makes

⁸ For a related problem, which is especially apparent when providing an outward-directed approach to intentions, see Way (2007).

sense to ask whether one should be in a perceptual state. To take a simple example, when experiencing the Müller-Lyer illusion, the mindreading system will inaccurately self-attribute the perception that the lines are equal in length.

Given these various problems, perhaps Carruthers should simply avoid defending an outward-directed account. The ISA theory is certainly not wedded to them. But, as it stands, Carruthers is faced with a dilemma: in holding that the mindreading system looks inwards to input from the mind, he must explain how privileged access to propositional states can be preserved despite the fleeting nature of sensory states; but, in avoiding this problem and holding that the mindreading system looks outwards to input from the world, he must meet the evidential objection and the limited application of such an approach.

5. Conclusion

Carruthers has integrated a wide range of empirical evidence to support a bold theory, and his overarching argument is impressive. However, we submit that his Interpretive Sensory-Access theory cannot easily maintain that self-attributions of attitudes are more reliable than our beliefs about other people's mental states (privileged access). Greater sensory information can lead to unreliability and sensory states are arguably more variable and fleeting than Carruthers recognizes. The ISA theory could avoid a more revisionary tack by treating sensory access as outward-directed, since then the mindreading system would avoid looking "inward" for its sensory input. But this strategy seems to involve unreliable inferences or transitions, especially when applied to perceptual or sensory states.

We are unsure whether Carruthers himself would simply welcome this result. He sometimes seems to avoid treating self-knowledge as no more reliable than other-knowledge, despite sharing the same basic mechanism. However, even if he were to embrace a more revisionary outlook, as he sometimes intimates, we believe this is *prima facie* an undesirable consequence of the ISA theory, as it denies the privileged access we seem to have to our own attitudes.⁹

⁹ Thanks to Jakob Hohwy, Jonathan Way, and anonymous referees for comments on a draft of this paper.

References

- Boyle, M. (2011). II—Transparent Self-Knowledge. *Aristotelian Society Supplementary Volume*, 85(1), 223-241.
- Byrne, A. (2011). I—Transparency, Belief, Intention. *Aristotelian Society Supplementary Volume*, 85(1), 201-221.
- Byrne, A. (2012). Review: The Opacity of Mind: An Integrative Theory of Self-Knowledge. *Notre Dame Philosophical Reviews*.
- Carruthers, P. (2010). Introspection: Divided and Partly Eliminated. *Philosophy and Phenomenological Research*, 80(1), 76-111.
- Carruthers, P. (2011). *The Opacity of Mind*: Oxford University Press.
- Evans, G. (1982). *The Varieties of Reference*: Oxford University Press.
- Hohwy, J. (2011). Phenomenal Variability and Introspective Reliability. *Mind & Language*, 26(3), 261-286.
- Hurlburt, Russell T. & Schwitzgebel, Eric (2007). *Describing Inner Experience? Proponent Meets Skeptic*. MIT Press.
- Owens, David (2003). Knowing Your Own Mind. *Dialogue* 42(4): 791–98.
- Ryle, Gilbert (1949/2009). *The Concept of Mind*. Routledge. (First published 1949 by Hutchinson.)
- Schwitzgebel, E. (2008). The Unreliability of Naive Introspection. *Philosophical Review*, 117(2), 245-273.
- Way, Jonathan (2007). Self-Knowledge and the Limits of Transparency. *Analysis* 67: 223-30.
- Zimmerman, Aaron (2008). Self-Knowledge: Rationalism vs. Empiricism. *Philosophy Compass* 3(2): 325–352.