



PROJECT MUSE®

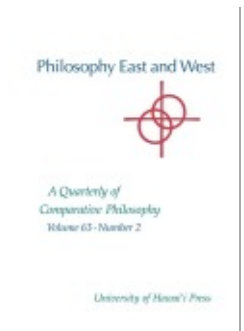
---

## Self-awareness: Eliminating the Myth of the “Invisible Subject”

Monima Chadha

Philosophy East and West, Volume 61, Number 3, July 2011, pp. 453-467  
(Article)

Published by University of Hawai'i Press  
DOI: [10.1353/pew.2011.0041](https://doi.org/10.1353/pew.2011.0041)



➔ For additional information about this article  
<http://muse.jhu.edu/journals/pew/summary/v061/61.3.chadha.html>

# SELF-AWARENESS: ELIMINATING THE MYTH OF THE “INVISIBLE SUBJECT”

**Monima Chadha**

School of Philosophical, Historical and International Studies, Monash University, Australia

## *Introduction*

In the sixth century A.D., in a debate with the Buddhists about the nature of Self, the well-known Naiyāyika Uddyotakara declared that there is no need prove that the Self or what is referred to by the pronoun “I” exists, for on that score there cannot be any significant disagreement.<sup>1</sup> It is only this or that specific metaphysical nature of the self that is the subject of controversy. To limit the scope of the debate at issue here, we employ the same strategy. It is beyond doubt that many cognitive processes involve consciousness of the self, for example monitoring one’s activities, as in learning how to dance or planning an important event in the near future. We can safely assume that there cannot be any significant disagreement about the fact that creatures like us are self-aware. The subject of our controversy is: what is one conscious of when one is conscious of oneself? Alternatively, which “self” is one conscious of when one is conscious of oneself? At this stage, I can only offer an incomplete answer. Self-awareness, whatever be the nature of it, cannot reveal the invisible subject. By the term “invisible subject” I mean a purely mental, non-extended, Cartesian *Res Cogitans*, the thinking subject that can never be an object of any conscious experience. David Hume’s famous denial, “when I enter most intimately into what I call myself . . . I can never catch myself at any time,”<sup>2</sup> is aimed at such an “invisible subject.” The invisible subject or pure Self, if it does exist, cannot be known by the ordinary methods of knowledge available to creatures like you and me. But there are other aspects of the self that are readily made available to us in reflection and in first-order reflexive acts like sensing pain. But, before we proceed with what has become a controversial issue about the nature of the self that is available to ordinary, finite beings like us, it is important to address some preliminary issues. First, we must define what we mean by self-awareness. And second, what are the various ways in which we may be self-aware.

In the first section in this essay, I will briefly describe what we mean by self-awareness and also give a brief outline of the accounts of self-awareness developed in the Indian and Western traditions. Not surprisingly, there are similarities in the theories constructed and defended in the Indian and Western traditions, and they fall under two broad categories: the *paraprakāśa* (literally, other-illumination) or reflectionist theories, and *svaprakāśa* (literally, self-illumination) or reflexivist theories.<sup>3</sup> I will only briefly describe these theories in the second section and will not offer arguments in favor of or against these views.<sup>4</sup> At first sight these two categories of theories seem to present competing accounts of self-awareness, and this is certainly how the

discussion is presented in the ancient Indian literature and in contemporary philosophy of mind. The two theories compete insofar as we want to answer the question as to what is the form of basic or fundamental self-awareness. In the third section, I argue against Chakrabarti that reflexivism is neither incoherent nor deeply incompatible with reflectionism. In the fourth section, I show that the reflexivist and the reflectionist accounts are not deeply incompatible, as the self can be both a subject and an object of experience, even though the awareness of the self as a subject is more basic and precedes the awareness of the awareness of the self as an object. The aim of the fifth section is to show that reflexivism does indeed have an upper hand in this debate since it alone characterizes the basic or fundamental form of self-awareness, pre-reflective awareness. Along the way, I address the controversial issue of which self is revealed by self-awareness.

In the last section, I propose that there is room for a pluralistic approach, rather than a monolithic approach that suggests an 'essential characterization' of the 'self' that is revealed in self-awareness. There is an over-abundance of delineations between various selves in the philosophical, psychological, cognitive science and psychoanalytical literature. It would be impossible to review all of these diverse notions of self in the course of this essay. So I will restrict myself to a distinction that has attracted the attention of philosophers in the Indian and Western traditions: is the self that we are conscious of when we are conscious of ourselves essentially an invisible subject? We turn to the preliminaries before discussing this question.

### *Preliminaries*

Following Roy Perrett, we can define the self-awareness thesis as the claim that if a subject is aware of an object, then the subject is also aware of being aware of that object.<sup>5</sup> The self-awareness thesis does not entail that the subject is explicitly aware of her awareness, nor does it entail that if the subject is in a mental state she knows that she is in that state. Thus, the self-awareness thesis is compatible with Freudian 'unconscious' states of which we are only implicitly aware. Also, the self-awareness thesis is much weaker than Descartes' Infallibility thesis, which implies that a subject cannot have mistaken beliefs about her own conscious states. It is important also to note that the self-awareness thesis does not entail that the self or the subject must be a relatively permanent center of consciousness. In fact, one version of the reflexivist theory offered by the Yogācāra Buddhist interprets the word "self" in "self-awareness" as awareness, not the subject of the awareness. In effect, an account of self-awareness is just meant to be an account of consciousness that explains how certain mental goings-on are illuminated. I will now turn to discuss the two broad categories of theories that attempt to offer an account of self-awareness.

Reflectionism can be described simply as a higher-order theory of self-awareness; that is, self-awareness involves a second-order awareness taking a distinct first-order awareness (say a perception) as its intentional object. According to reflectionism, awareness of a mental state or cognition is *paraprakāśa*; that is to say, it is illuminated or revealed by another awareness. There are two ways to develop the idea depending

on how we understand the notion of second-order awareness. Perhaps we become aware of seeing a cup because we perceive, with some inner eye, that we are seeing a cup. Or perhaps we just become aware of seeing a cup, because we come to believe that we are seeing a cup. Accordingly, the second-order awareness consists in either perceiving or believing that one is in a certain perceptual or some other mental state. The Naiyāyikas, John Locke, and David Armstrong favor the first option: self-awareness is no more than perception of inner mental states by the subject. This class of reflectionist theories falls under introspectionism. The second option, that of construing self-awareness as belief, also known as the higher-order-thought (HOT) hypothesis, is espoused and defended most famously by David Rosenthal.<sup>6</sup> In the present article, I will ignore the HOT hypothesis because it aims to offer a reductive account of consciousness, an enterprise that, as various critics rightly point out, is doomed to failure.<sup>7</sup> Unlike the HOT hypothesis, the introspectionists do not aim to give a reductive account of all kinds of consciousness in terms of higher-order awareness. Indeed, the Naiyāyikas claim that consciousness is not at all explained by second-order awareness or introspection.<sup>8</sup> Most versions of introspectionism leave open the possibility that the subject of a mental state is not always introspectively aware of being in that state. Does this mean that the subject is not aware at all of being in that mental state? This is a question to which we will return in the next section.

In contrast, reflexivism holds that an awareness of an object reveals itself at the same time as it reveals an object. In this sense awarenesses are *svaprakāśa* or self-illuminating. Like reflectionism, there are multiple versions of reflexivism, but they all agree that the same-order theory is the best explanation of basic self-awareness. In the strongest version, held by the Yogācāra Buddhists, awareness is only ever aware of itself. Since the Buddhists deny the existence of the self, there is no subject, over and above or behind the stream of experience, to view the passing show of experience. The self is not acquainted with the experience; rather, experience is acquainted with itself. The Yogācārin holds that the intentional object of awareness is not independent of the awareness. But some more moderate versions of reflexivism defended by the Prabhākara Mīmāṃsākas in the Indian tradition hold that a mental awareness reveals its intentional object to itself and to the Self. In the Western tradition, historically, the divide between reflectionism and reflexivism coincides with the divide between representational and phenomenological approaches. However, recently, some representationalists have put forward accounts of self-awareness that are clearly reflexivist.<sup>9</sup> Phenomenologists hold a version of reflexivism according to which the intentional aspect is directed toward the object in the world and the reflexive aspect consists of a basic, pre-reflective awareness of the self as the subject of the awareness. Representationalists, who favor the same-order account of self-awareness, talk about a complex conscious state that is formed by a fusion of two separate vehicles, one of which is about the world and the other about the first state.

The foregoing discussion will suffice to give the reader a basic idea of what we mean by self-awareness and the reflectionist and reflexivist accounts of self-awareness. Now we proceed to argue for the substantive and controversial claims that I hope to defend in this essay.

## Reflectionism and Reflexivism

Let us begin by noting that reflexivists do not claim that self-awareness is not exclusively being reflexively aware of a mental state, for we can be reflectively self-aware. They claim that being aware of an object presupposes being reflexively aware of being aware. In normal cases of perceiving—say, when I see a cup—I am aware of the cup and pre-reflectively aware of something like what it is for me to see a cup. However, it is possible for me to turn my attention from the cup to the awareness of the cup and thus be reflectively aware of my awareness. On the other hand, most reflectionists do not claim that self-awareness solely consists in being reflectively aware of a first-order mental state. A notable exception is David Rosenthal, who defends the HOT hypothesis, which offers a reductive account of consciousness or self-awareness in terms of higher-order thoughts. However, as I said earlier, we will ignore the HOT hypothesis for our purposes here.

Introspectionists, for example David Armstrong, distinguish between several types of consciousness: minimal consciousness (e.g., faint sensation), perceptual consciousness (e.g., awareness of what is going on in the body and one's immediate environment), and introspective consciousness (e.g., awareness of the current state and activities in one's own mind).<sup>10</sup> Introspective awareness involves reflective awareness of one's first-order mental states. But surely perceptual consciousness in Armstrong's sense includes self-awareness, even if it is awareness of one's bodily self. The Naiyāyikas also do not think that the higher-order awareness is necessary for consciousness; for them, first-order states (*vyavasayas*) are conscious whether or not they are perceived. The Naiyāyikas hold that mental states are not automatically and reflexively apperceived; rather, introspective awareness or apperception (*anuvyavasya*) requires focusing attention on one's own mental states. So, to return to the question raised above, we can say that it is possible for a subject to be aware of being in a mental state, without being introspectively aware of being in that mental state. Thus, it seems that neither the reflexivists nor the introspectionists claim that self-awareness is wholly and solely constituted by reflexive awareness or reflective awareness, respectively. So, what is the disagreement about?

The debate between the reflectionist and the reflexivist reveals disagreement on many points, but I think it would be fair to say that there are two main points of disagreement that deserve discussion. First, the introspectionist denies that we can ever be reflexively aware of our mental states, because reflexive awareness collapses the distinction between the cause of the awareness and the awareness itself. I will deal with this objection in the next section. We will also briefly consider the charge of incoherence against reflexivism toward the end of the third section. Second, on the other hand, the reflexivist complains that the introspectionist view reduces the essential subject to a mere thing. Insofar as the causal-perceptual model of awareness turns another mental state and/or the subject into an object of experience, it fails to do justice to the fundamental or basic character of consciousness. The disagreement between the reflexivist and reflectionist is about the form of basic or fundamental self-awareness. The reflexivist considers pre-reflective self-awareness as a type of

self-awareness that precedes and is more basic than reflective self-awareness. I shall explore these issues in the fourth section to show that the reflexivist and the reflectionist accounts are not deeply incompatible, since the self can be both a subject and an object of experience, even though the awareness of the self as a subject is more basic and precedes the awareness of the awareness of the self as an object. In the fifth section the discussion will reveal that the reflexivist is right insofar as we cannot avoid the regress problem without accepting something like basic pre-reflective awareness that is non-objectifying and non-criterial self-awareness. Moreover, pre-reflective self-awareness is understood as a type of self-awareness that emphatically lacks any kind of dyadic structure and relational process of self-objectification.

### *Why not Reflexivism?*

The Naiyāyikas subscribe to a largely causal-perceptual model of awareness, so it is not surprising that they, *pace* Armstrong, reject reflexive awareness on the basis of the so-called distinct existences argument. Udayana's version of the distinct existences argument starts with the commonsense realist premise that our experience reveals that the apprehensible object is distinct from the awareness of the object.<sup>11</sup> The reflexivist who wants to deny this commonsense assumption will have to argue for a thesis of universal delusion to establish that an ordinary perception that registers a difference between the awareness and the object of awareness is always mistaken. In dreams or double-moon visions our awareness registers a duality that can be shown to be false by means of an argument. But in ordinary perception no such argument is forthcoming to reveal the unity rather than the distinctness of an object and its awareness. And so the Naiyāyikas argue that if we cannot establish nonduality between the apprehensible object and its awareness for awarenesses in general, there is no reason to hold that there is nonduality in the case of apperception or introspective awareness of mental states. Armstrong's version of the argument establishes the distinctness between the introspective awareness of a mental state and the mental state itself by pointing out that a mental state is part of the causal complex responsible for its apperception or introspective awareness.<sup>12</sup> The Naiyāyikas would agree with this starting point. Furthermore, since cause and effect must be distinct existences, no event, and thus no mental event, can cause itself. Since a mental state is partly causally responsible for its awareness, it follows that mental state is distinct from our awareness of that state.

Even if we accept that Udayana's claim that our ordinary perception registers a difference between the apprehensible object and its awareness, we need a further argument as to why this should be so in the case of apperception. Armstrong's argument does not do any better; even if we accept that cause and effect are distinct existences, we need a further argument to show that the relation between a mental state and its awareness is a causal relation. The Buddhist philosophers Dinnāga and Dharmakīrti appeal to the sensations of pain, pleasure, anger, et cetera, as paradigmatic examples of reflexively aware cognitive states. The standard example of pain is used to emphasize the double character of awareness, which reveals the object as-

pect as well as the awareness aspect. To highlight the idea that cognitive states are essentially self-aware, Dharmakīrti asks a rhetorical question: how can one say, “My head aches, but I am not aware of it”? Thus, he argues, a cognition must be aware of itself as soon as it arises.<sup>13</sup>

Udayana counters the Buddhist claim by arguing that pain, pleasure, anger, et cetera are cognized as soon as they arise only because of their characteristic intensity.<sup>14</sup> But other feeble cognitions may pass by completely unnoticed. Or, if the subject is too distracted or too sleepy to look back at what she just cognized, the subject may not be self-aware. The Naiyāyikas argue that it is only a contingent fact that an awareness is cognized as soon as it arises. Śāntaraksita responds to this Nyāya argument by claiming that if a cognition is not self-aware, it loses its essential nature, illuminatorship (*prakāśata*). Unconscious, inert objects depend on other objects for their illumination. If an awareness, which is a form of consciousness, becomes dependent for its illumination, then there would be no distinction between conscious and unconscious matter.<sup>15</sup> The Nyāya rejects this argument by saying that the essential nature of cognition is not self-illumination, but illumination of its objects. And, since unconscious matter does not have the property of illuminating other objects, it can be easily distinguished from consciousness. The arguments offered by the Buddhist do not suffice to persuade the Naiyāyikas to embrace the thought that we are reflexively aware of some of our mental states.

Indeed, Arindam Chakrabarti, a contemporary proponent of the Nyāya position, has recently formulated a new argument against reflexivism.<sup>16</sup> The argument has the form of a *reductio* that assumes reflexivism in the hope of deriving contradiction to conclude that reflexivism is false. Chakrabarti begins by stating the reflexivist thesis that a cognition of an object is constituted by another cognition of the cognition of the object. Thus, he claims, “Cognition of a blue pot (Cbp) = Cognition of cognition of blue pot (Ccbp) or cognition that one is aware of a blue pot.” Furthermore, Chakrabarti claims that the cognition that one is aware of a blue pot is not the same as the cognition that one is aware of a pot. The introduction of an additional qualifier to the content of the cognition changes, even if the external object of the cognition does not change. Thus, he concludes that the cognition of a blue pot is not identical with the cognition of a pot. That is to say, “Cbp is not identical with Cp.” At this point Chakrabarti appeals to simple realist wisdom to assert that a qualified object is none other than the sheer object; a blue pot is a pot nonetheless. This simple claim, in turn, is exploited to draw a surprising consequence. Chakrabarti writes:

So there is a pot such that whoever is *cognizing* it as a sheer pot is cognizing the blue pot and whoever is cognizing the blue pot is also cognizing a pot whether they recognize this identity or not. So at least in this one case, cognition of bp is identical with the cognition of p, thus Cbp is identical with Cp [*italics added*].<sup>17</sup>

It is hard to understand what Chakrabarti is asserting because there are different ways of interpreting the term “cognizing” in the quote above. I will focus on the second occurrence of “*cognizing*” to explicate the two interpretations.

In the first interpretation, the occurrence of “cognizing” in “cognizing the blue pot” may be shorthand for a non-simple awareness that has a qualificand-qualifier structure.<sup>18</sup> Thus, the first conjunct reads: there is a pot such that whoever is cognizing it as a sheer pot is cognizing it as a blue pot. But in this reading the two cognitions are not identical. So, the first conjunct is simply false. It is worth noting here that this first is the obvious interpretation, as Matilal points out that perceptual awareness is non-simple, though it can become simple in exceptional cases and only under some theoretical consideration.<sup>19</sup> The same point is sometimes made by claims such as: all seeing is seeing-as.<sup>20</sup> In the second interpretation, the “cognizing” in “cognizing a blue pot” signifies a simple, unanalyzed awareness that lacks the qualificand-qualifier structure. Even though this interpretation is not natural, a simple awareness can be abstracted from a non-simple awareness. According to the second interpretation the two occurrences of “cognizing” in the first conjunct signify two different kinds of cognitions or awarenesses, and thus the identity claim is false again. The identity between the cognition of a blue pot and the cognition of a pot is required to derive a contradiction, and since the identity claim is false, Chakrabarti’s argument against reflexivism fails. Reflexivism is not incoherent.

### *Why not Introspectionism?*

The Buddhist may attempt to repudiate the claim that self-awareness requires introspective access to one’s own mental states. This brings us to the second point of difference mentioned above. Some contemporary philosophers deny that there is any need for a second-order representation or introspectionist mechanism to explain self-awareness of perceptual states. This view is indeed held by some contemporary reflexivists. S. Shoemaker suggests that the introspectionist view is wildly implausible:

No one thinks that in being aware of a sensation or sensory experience one has yet another sensation that is “of” the first one, and constitutes its appearing to one a certain way. . . . and no one thinks that there is such a thing as an introspective sense-experience of oneself, an introspective appearance of oneself that relates to one’s beliefs about oneself as visual experiences of things one sees relate to one’s beliefs about those things.<sup>21</sup>

The reason one is not presented to oneself “as an object” in self-awareness is that self-awareness is not perceptual awareness, i.e. not a sort of awareness in which objects are presented.<sup>22</sup>

This strong denial of introspectionism is endorsed by reflexivists. They argue that the perceptual model of self-awareness is inappropriate because it fails to take into consideration the essentially subjective nature of consciousness. Since higher-order awareness takes another state or subject as its object, it reduces the subject to a mere thing. But the reflexivists insist that the subject is precisely that which can never become an object. We are directly aware of the self in the reflexive mode of awareness, which is a non-dyadic awareness in the sense that it does not involve the subject-



object structure. Recently, Arindam Chakrabarti questioned a key premise in this argument, namely that the subject is precisely that which can never become an object.

Chakrabarti targets what he calls the biggest conceptual difficulty that philosophers have voiced about introspective self-awareness: that the very idea of making an object out of a subject is incoherent.<sup>23</sup> Chakrabarti begins by pointing out that neither linguistic-conceptual intuitions nor phenomenological intuitions can guide us here, so he sticks his neck out for a positive argument for his thesis that “self-objectifying self-awareness must be possible, whatever may be the nature of that self which perceives itself, and whatever the nature of that perception.” His argument is as follows:

P1: If it is impossible to be directly acquainted with the referent of *A*, then I cannot perceive that, wonder whether, mistakenly feel that, or come to realize that I did not notice that *A* is *f*, when “*A*” is a directly referential singular term.

P2: I can perceive that, wonder whether, mistakenly feel that, or come to realize that I did not notice that I am jealous or I am pleased.

P3: “*I*” is a directly referential singular term, albeit one which has some unique features and constraints.

Therefore, it must be possible for me to be directly acquainted with the referent of “*I*.”

Chakrabarti is quick to notice that premise 3 denies G.E.M. Anscombe’s famous thesis that the pronoun “*I*” has no reference at all.<sup>24</sup> Anscombe’s argument rests on the claim that a word can genuinely refer to an object only if it is possible for us to mis-identify the referent while using that word. Since it is impossible to commit that kind of identificatory mistake with “*I*,” Anscombe argues that the term “*I*” lacks a reference.

Chakrabarti tries to argue that an identificatory mistake is possible by offering a rather unhelpful example from Martin and Hogan.<sup>25</sup> I think Chakrabarti’s discussion is focusing on the wrong premise. Semantic intuitions are just as unhelpful as linguistic or phenomenological intuitions. For, as Chakrabarti himself points out, Bertrand Russell’s intuition directly contradicts Anscombe. Russell argued that the only genuine proper name or reference guaranteed a singular term is the pure demonstrative “*this*,” which picks out, so to say, a featureless pure subject.<sup>26</sup> Surely, there is no error regarding the bare particular whether it is picked out by a “*this*” or an “*I*,” but for that very reason the bare particular eludes our grasp, too.

The debate about the reference of “*I*” is beside the point here. Our concern is whether as an object of introspection the self has a status similar to the status of other objects of other kinds of perception. In cases of genuine perception and perceptual judgments the possibility of error arises only with respect to the content of the concepts or predicates that we ascribe to the object of perception. Peter Strawson is absolutely right in pointing out that there is no pure seeing: all seeing is seeing-as.<sup>27</sup> I see this thing on the table in front of me as a computer. The question is whether the object or the self as qualified by properties or having features is immune to error

through mis-identification. And the obvious answer is: No! I may mistakenly think I am angry when I am jealous, or I may feel that I need to smoke when in fact I am hungry. The self is vulnerable to mistakes and partial occlusion, and this is what is required for it to be a perceptual object in the broad sense. The self as qualified by properties or features is like any other perceptual object, and thus there seems to be no *prima facie* reason to regard it as essentially a subject that is somehow violated if it is sometimes contemplated as an object of inner perception or introspection. This argument shows that there is no logical or conceptual hurdle in accepting that the self is available in experience as an object qualified by properties.

Furthermore, it is important to note that perhaps Chakrabarti is being unfair to the reflexivist. The reflexivist never makes the strong claim that the subject can never be made into an object of experience. Rather, the reflexivist claim is that one can in fact succeed in taking oneself as an object, but in doing so one is presupposing a prior non-objectifying, non-criterial self-awareness as the condition of the possibility of self-objectification.<sup>28</sup> Dan Zahavi argues for this weaker claim by pointing out that the object-use of “I,” as opposed to the subject-use of “I,” as in “I am bleeding,” is typically used for ascribing a certain property to a body. Nevertheless, the object-use of “I” does qualify as self-ascription because we are talking about my body and not any body. And, we cannot explain what is meant by calling a body “my body” without referring to the subject-use of “I,” as in saying that my body is the body with which I speak, hear, act, et cetera. Thus, the object-use of “I” and more generally every objectifying self-ascription piggybacks on a prior non-objectifying self-reference. The point is that since the self can be an object of experience it is neither essentially nor exclusively a subject.<sup>29</sup>

Before we go any further let us take stock of the situation. In this section, my aim has been to show that the reflexivist and the introspectionist accounts of self-awareness are not deeply incompatible. We have seen that there is no logical or conceptual incoherence in accepting that the self can be both the subject and object of introspection. Thus, the reflexivist’s worry that introspectionism reduces the essential subject to a mere object is baseless, for the Self is not essentially a subject. The self can be both an object and subject of perception.

### *Reflectionism and Reflexivism Revisited*

We have seen that introspectionists accept that some forms of consciousness, for example perceptual consciousness, which includes awareness of one’s body and the immediate environment around oneself, are not introspective. Is there any reason to think that such forms of consciousness are best explained by the reflexivist’s account of self-awareness? This is the next task. Perhaps the best way to persuade the Naiyāyikas and other introspectionists to accept that we are reflexively aware of some mental states is to offer a version of the regress argument that was indeed applied to the Nyāya view by the reflexivists in the Indian tradition. Recall that the introspectionists, Naiyāyikas, and Armstrong attempt to explain what makes a state self-aware in terms of its being the object of a higher-order perception. Now either

the higher-order perception is self-aware or it is not. If it is itself self-aware, then, according to the introspectionist, it will require a third-order representation. Furthermore, if the third-order representation is self-aware, we are off on a vicious regress. On the other hand, if the higher-order perception is not self-aware, then the introspectionist owes us an account of how this representation reveals its object, in this case a first-order awareness, to the subject. Any response given by introspectionists, or reflectionists more generally, seems either ad hoc or plainly false. The introspectionist needs to explain how a mental state can reveal its object without revealing itself. The Nyāya response that some cognitions are not self-aware and thus manage to slip away unnoticed will not work in this instance, since the subject is aware of the object of the cognition or the representation, but not self-aware of the cognition or representation itself.

Furthermore, several critiques of introspectionism have decisively argued that higher-order theories cannot explain phenomenal consciousness. A mental state that is phenomenally conscious is a state in which there is something to be like. It is the sort of state one is in when one is in pain, or enjoying a fine piece of music, or tasting a rather plain pasta dish. In a recent article, Josh Weisberg spells out the problem succinctly by alluding to the problem of misrepresentation.<sup>30</sup> Higher-order theorists suggest that a distinct higher-order representation makes us conscious of our conscious states. But there is always the possibility of misrepresentation; the higher-order state may not correctly represent the first-order state, or a more radical failure may result in cases when we have a higher-order representation that is targetless. Since the higher-order representation determines what it is like for the subject, a case of misrepresentation is subjectively indistinguishable from the veridical case. Thus, it is immaterial which first-order state, if any, is represented by the higher-order representation. This renders the relational structure of higher-order theories explanatorily superfluous.

Insofar as the reflexivist theories do not posit two distinct representations to explain self-awareness or consciousness, they are potentially better suited to avoid both of the pitfalls raised above. Moreover it seems obvious in perceptual cases that we are simultaneously aware of the properties of the object as well as the representation. An example from Mathew Mackenzie makes the point clearly:

Suppose that I have a perception, for example of a cup in the room, of which I am completely unaware. Because the mental state reveals nothing about itself I will not be aware whether the perception is visual or tactile, since being visual (or tactile) is a property of the representation, not its object. But conscious perception isn't like that at all. When I have a conscious visual perception of a cup, I am aware of the cup as it is presented to me in vision—indeed, I am aware of what the cup looks like. . . . This suggests that I am aware of the properties of the cup and properties of my representation at the same time, thus allowing me to be aware that I am seeing the cup rather than touching just by having that perception.<sup>31</sup>

This example functions as an intuition pump to woo the higher-order theorist to take the reflexivist account of self-awareness seriously. But if the reflexivist wishes to win over the higher-order theorist, the former must show that reflexivism can avoid the

regress problem and offer an adequate explanation of phenomenal consciousness. The answer to the regress problem is straightforward and does not require much argument. The regress can be blocked simply by insisting that some representations are reflexively self-aware.

What about the second worry: can reflexivists give a satisfactory explanation of phenomenal consciousness? Several attempts in the literature have been made to show how the reflexivist might address this problem. I will discuss a recent model proposed by Robert Van Gulick because it combines the virtues of the representationalist and phenomenological same-order theories and thus has the potential to offer a satisfactory explanation of phenomenal consciousness, if it can indeed be explained.<sup>32</sup> He labels the approach as “Higher-order Global States” (HOGS). The idea is to deny the distinctness between the higher-order state and the object state and treat the transition from unconscious state to conscious one as a matter of “recruiting it into a globally integrated complex whose organization and intentional content embodies a heightened degree of reflexive self-awareness.”<sup>33</sup> The model explicates conscious content as implicit in the phenomenal structure of the conscious state rather than as explicitly borne by a distinct higher-order state. But note that the model takes the phenomenal structure as primitive rather than offering an explanation. Van Gulick acknowledges this:

I do not suppose that the HOGS model will in itself explain everything we want to understand about phenomenal consciousness; . . . I doubt any model could suffice for complete understanding. Nor is the HOGS model intended as a reductive theory; it aims to provide some insights into what phenomenal consciousness is and how it could be realized that might otherwise be missed. And if it does so, that might be enough to count as at least a partial explanatory success.<sup>34</sup>

This sentiment is echoed by other reflexivists in the analytic as well as the phenomenological traditions. As far as a satisfactory explanation of phenomenal consciousness is concerned, the most that reflexivism can achieve is to show that the phenomenal consciousness is partially, though non-reductively, explained in terms of reflexive self-awareness. The model explicates conscious content as implicit in the phenomenal structure of the conscious state rather than as explicitly borne by a distinct higher-order state. Van Gulick’s model does not aim to offer a reductive explanation of phenomenal consciousness; rather it takes the notion of implicit self-awareness as primitive. This sounds very similar to Sartre’s idea of pre-reflective awareness, but he fails to acknowledge this.

The reason why analytic philosophers shy away from acknowledging the notion of pre-reflective reflexive awareness is because it is hard to cash out the notion in a way that it can do the real work of philosophical analysis. This is precisely because pre-reflective awareness does not correspond to any explicit, propositional/conceptual, self-conscious experience of the self as an object or mere aggregate of objects. The only way, I think, that we can make sense of the notion of the pre-reflective self is to treat it as a bodily subject. Merleau-Ponty endorses this view in rejecting the idea that the subject is a separate transcendental ego that perceives and interprets sensation closed up behind a world of ideas. Merleau-Ponty coins a new

notion of body-subject as that which sets the conditions for experience. The body-subject is neither an external object (e.g., it is not my empirical body that I can observe in the mirror) nor an internal object. When I am conscious of myself as the subject of an experience, I am not scrutinizing an internal self looking at the external world; I am simply looking outside at the external world, and within this single act of consciousness I pre-reflectively experience myself-as-subject. Merleau-Ponty puts the point thus: The consciousness of the world and pre-reflective self-consciousness “are strictly contemporary. There is a world for me because I am not unaware of myself; and I am not concealed from myself because I have a world.”<sup>35</sup> The body-subject does not simply undergo the influence of the world in an adaptive or passive way, but rather actively orients and reorients itself while it undergoes various worldly influences. The body-subject actively undergoes and negotiates from within the world a perspectival experience oriented in a particular way.

Contemporary phenomenologists Dan Zahavi, S. Gallagher, and Dorothee Legrand explain the notion of pre-reflective awareness by using something like Merleau-Ponty’s notion of body-subject.<sup>36</sup> Legrand, for example, characterizes the pre-reflective self-consciousness as the “performative” body and the “transparent” body. The performative body is perhaps best exemplified by the pre-reflective experience of an expert dancer who knows his choreography and improvises a skillful dance. In this case, observational consciousness of the body is not in any way necessary in order to control the body’s movements, and may well be counterproductive. But normal people, and not just expert dancers, are also pre-reflectively aware of the performative body in everyday basic intentional actions. We project ourselves into the world through our body and actions, but are normally only conscious of the goals rather than the means to achieve them. Thus, explicit consciousness of and attention to the outside world tend to hide the pre-reflective experience of the performative body, although it is present. Legrand describes the transparent body thus: “The body is transparent in the sense that one looks through it to the world.”<sup>37</sup> The pre-reflective bodily experience is precisely the experience of the world as given through the transparent body. The latter is not perceived as an object but experienced specifically as a subject perceiving and acting in the world.

The following quote from Gareth Evans also very aptly describes the notion of the transparent body:

[W]e are able to know our position, orientation, and relation to other objects in the world upon the basis of our perceptions of the world. Included here are such things as: knowing that one is in one’s bedroom by perceiving and recognizing the room and its contents; knowing that one is moving in a train by seeing the world slide by; knowing that there is a tree in front of one, or to the right or left, by seeing it; and so on.<sup>38</sup>

In this way, we are explicitly conscious of objects in the outside world through the simultaneously performative and transparent intermediary of the living human body. Neither of these forms of pre-reflective self-consciousness takes the self to be the object of experience, but instead the pre-reflective self is the essentially embodied subject of experience.

That said, I don't think that the notion of pre-reflective awareness as explained here in terms of bodily awareness is sufficient for satisfactory explanation of phenomenal consciousness. At most, it provides the primitive ground for phenomenal consciousness. So, we can conclude that phenomenal consciousness can be partially, though non-reductively, explained by the reflexivist model. This does not give a decisive victory to reflexivist accounts of self-awareness. But it is clear, I hope, that a theory of phenomenal consciousness that does not take ubiquitous primitive, reflexive self-awareness into account is a non-starter. So, the reflectionist cannot hope to do better without embracing the reflexivist model as part of their account of self-awareness.

### *Conclusion*

I have shown that the reflectionist and reflexivist theories of self-awareness are not deeply incompatible; rather they present or reveal different aspects of the self. Introspectionism, I claim, gives an insight into the self as an object that figures in our mental states and experiences. The urge to know the self as a subject is satisfied in our reflexive experiences. However, we realize that it is not the invisible subject that we were hoping to meet. For as we saw, the best way to make sense of pre-reflective awareness is to think of it as the awareness of a bodily subject. In conclusion, the introspectionism and the reflexivism together dispel the myth of the "invisible subject." The self is neither essentially a subject, for it can be an object of perception, nor is it essentially invisible, for it is at least a bodily subject that we come across in the act of seeing and acting in ordinary, everyday situations.

### Notes

*I am very grateful to an anonymous referee for this journal for very useful comments. This article has been greatly improved by addressing his specific concerns.*

- 1 – Arindam Chakrabarti, "The Nyaya Proofs for the Existence of the Soul," *Journal of Indian Philosophy* 10 (1982): 211–238.
- 2 – David Hume, *A Treatise of Human Nature*, ed. L. A. Selby-Bigge, revised by P. H. Nidditch (1739–1740; Oxford: Clarendon Press, 1978), p. 252.
- 3 – The broad characterization of Western theories and their Indian counterparts under one umbrella has recently been put forward in an article by Mathew D. Mackenzie, "The Illumination of Consciousness: Approaches to Self-Awareness in the Indian and Western Traditions," *Philosophy East and West* 57 (2): 40–62.
- 4 – The interested reader is directed to the article in the preceding footnote by Mackenzie, who does an excellent job of reviewing some of the central arguments and their weaknesses.

- 5 – Roy W. Perrett, "Intentionality and Self-Awareness," *Ratio* 16 (3) (2003): 222–235.
- 6 – David Rosenthal, "Thinking that One Thinks," in *Consciousness: Psychological and Philosophical Essays*, ed. Martin Davies and Glyn W. Humphreys (Oxford: Basil Blackwell, 1993), pp. 197–223.
- 7 – See, for example, A. Byrne, "Some like it HOT: Consciousness and Higher-Order Thoughts," *Philosophical Studies* 86 (1997): 103–129.
- 8 – Mackenzie points out that the Nyāya account of apperception is not an account of consciousness per se, but only an account of introspective awareness. See Mackenzie, "The Illumination of Consciousness," p. 42.
- 9 – U. Kriegel, "The Same-Order Monitoring Theory of Consciousness," in U. Kriegel and K. Williford, eds., *Self-Representational Approaches to Consciousness* (Cambridge MA: MIT Press), pp. 143–170.
- 10 – David Armstrong, "What Is Consciousness?" in Ned Block, Owen Flanagan, and Güven Güzeldere, eds., *The Nature of Consciousness* (Cambridge, MA: MIT Press, 1968), pp. 721–729.
- 11 – Udayana, *Atmatattvaviveka*, p. 230.
- 12 – David Armstrong, *A Materialist Theory of the Mind* (London: Routledge and Kegan Paul, 1968), p. 106.
- 13 – Dharmakīrti, *Pramānavārttika, Pratyaksha* chapter, v. 328.
- 14 – Udayana, *Nyayavarttikataparya-parisuddhi* (Bibliotheca Indica), p. 121.
- 15 – Śāntaraksita, *Tattvasamgraha*, v. 2020.
- 16 – Arindam Chakrabarti "So What, If the Subject becomes an Object in Self-Perception?" (unpublished manuscript, presented at APA Central Division meeting in Chicago, 2005).
- 17 – Ibid.
- 18 – This distinction between simple and non-simple awareness is made by Matilal. See B. K. Matilal, *Perception: An Essay on Classical Indian Theories of Knowledge* (Oxford: Clarendon Press, 1986), pp. 180–181.
- 19 – Ibid., p. 180.
- 20 – Ibid., p. 181.
- 21 – S. Shoemaker, "Self-Knowledge and 'Inner Sense,' Lecture I, The Object Perception Model," *Philosophy and Phenomenological Research* 54 (2): 249–269.
- 22 – S. Shoemaker, "Personal Identity: A Materialist's Account," in S. Shoemaker and R. Swinburne, eds., *Personal Identity* (Oxford: Blackwell, 1984), pp. 67–132, at p. 105.

- 23 – Arindam Chakrabarti, “So, What, If the Subject Becomes the Object in Self-perception?” (unpublished manuscript).
- 24 – G.E.M. Anscombe, “The First Person,” in Samuel Guttenplan, ed., *Mind and Language* (Oxford: Clarendon Press, 1975), pp. 45–65.
- 25 – Melinda Hogan and Martin Raymond, “Introspective Misidentification: An I for an I,” in *Advances in Consciousness Research* 30 (2001), *Self-Reference and Self-Awareness*, ed. Andrew Brook and Richard C. DeVidi, pp. 205–213.
- 26 – Bertrand Russell, “The Philosophy of Logical Atomism,” reprinted in *Logic and Knowledge*, ed. R. C. Marsh (London: George Allen and Unwin, 1956).
- 27 – P. F. Strawson, “Imagination and Perception,” reprinted in *Freedom and Resentment, and Other Essays* (London: Methuen, 1974).
- 28 – Dan Zahavi, *Self-Awareness and Alterity: A Phenomenological Investigation* (Evanston, IL.: Northwestern University Press, 1999).
- 29 – I thank an anonymous referee for this journal who helped me clearly appreciate this point.
- 30 – Josh Weisberg, “Same Old, Same Old: The Same-order Representational Theory of Consciousness and the Division of Phenomenal Labor,” *Synthese* 160 (2) (2008): 161–181.
- 31 – Mackenzie, “The Illumination of Consciousness,” p. 58.
- 32 – Robert Van Gulick, “Mirror, Mirror—Is that All?” in Uriah Kriegel and Kenneth Williford, eds., *Self-Representational Approaches to Consciousness* (Cambridge, MA: MIT Press, 2006), pp. 11–39.
- 33 – *Ibid.*, p. 15.
- 34 – *Ibid.*, p. 38.
- 35 – *Ibid.*, p. 298.
- 36 – See Dan Zahavi, *Subjectivity and Selfhood: Investigating the First-Person Perspective* (Cambridge, MA: MIT Press, 2005), and S. Gallagher, *How the Body Shapes the Mind* (Oxford: Clarendon Press, 2005).
- 37 – Dorothee Legrand, “Pre-Reflective Self-Consciousness: On Being Bodily in the World,” *Janus Head* 9 (2) (2007): 493–519.
- 38 – Gareth Evans, *Varieties of Reference* (Oxford: Clarendon Press, 1982), p. 222.