

REVUE DE
MÉTAPHYSIQUE
— ET DE —
MORALE

Fondée en 1893 par Xavier Léon et Élie Halévy

<i>Directeur de la publication</i>	Didier Deleule
<i>Rédactrice en chef</i>	Isabelle Thomas-Fogiel
<i>Secrétaire de rédaction</i>	Alain de Seze
<i>Bureau</i>	Renaud Barbaras, Bernard Bourgeois, André Charrak, Didier Deleule, Catherine Kintzler, Isabelle Thomas-Fogiel
<i>Comité de rédaction</i>	Anne Baudart, André Charrak, Pascal David, Denis Kambouchner, Sandra Laugier, Christiane Menasseyre, Jean-Michel Muglioni, Pierre Osmo, Emmanuel Picavet
<i>Comité de lecture</i>	Frédéric Brahami, Roger Bruyeron, Anissa Castel-Bouchouchi, Hugues Choplin, Michel Fichant, Thierry Gontier, Simone Goyard-Fabre, Francis Jacques, Guillaume Le Blanc, Béatrice Longuenesse, Pierre-François Moreau, Alexis Philonenko, Alain Renaut, Elisabeth Schwartz, André Tosel, Anne Gabrièle Wersinger
<i>Rédaction</i>	PUF - Revue de Métaphysique et de Morale, 6, avenue Reille, 75014 Paris E-mail : revue-metaphysique@wanadoo.fr
<i>Ventes et abonnements</i>	E-mail : revues@puf.com Presses Universitaires de France, Département des Revues, 6, avenue Reille, 75014 Paris, France Tél. : +33 1 58 10 31 00 – Fax : +33 1 58 10 31 82 IBAN : FR40 2004 1000 0100 3923 3A02 062 BIC : PSSTFRPPPAR

	<i>Particulier</i>	<i>Institution</i>
<i>Abonnement annuel (2011) France :</i>	54,00 € TTC	67,00 € TTC
<i>TVA : 2,10 % Étranger :</i>	67,00 € HT	81,00 € HT

**Revue de Métaphysique
et de Morale**

OCTOBRE-DÉCEMBRE 2010 – N° 4

PUBLIÉE AVEC L'AIDE DU CENTRE NATIONAL DU LIVRE
ET SOUTENUE PAR L'INSTITUT DES SCIENCES HUMAINES ET SOCIALES DU CNRS

LE MOI/THE SELF/LE SOI

NUMERO DIRIGÉ PAR BÉATRICE LONGUENESSE

SOMMAIRE

Béatrice LONGUENESSE, Présentation – De différentes manières de se rapporter à soi	419
Rolf-Peter HORSTMANN, The Limited Significance of Self-Consciousness	435
John PERRY, Persons and Selves	455
François RÉCANATI, Le soi implicite	475
Vincent DESCOMBES, Le marteau, le maillet et le clou	495
Christopher PEACOCKE, Self-Consciousness	521
Quassim CASSAM, How We Know What We Think	553
<hr/>	
Recensions critiques	571

Self-Consciousness

ABSTRACT. — *I distinguish two varieties of self-consciousness. One variety I label 'perspectival self-consciousness'. I propose an account of its nature, and consider its relations to: Gallup's mirror test for self-consciousness; Shoemaker's conception of immunity to error through misidentification; the possession of a conception of many minds; and some of Sartre's ideas on what it is to conceive of oneself as an object. A second variety of self-consciousness I label 'reflective self-consciousness'. I offer an account of this, its epistemological significance, and consider the ways in which perspectival self-consciousness and reflective self-consciousness have to cooperate if a thinker is to attain certain epistemic goals. I conclude with some reflections on the bearing of the metaphysics of subjects of consciousness, and the metaphysics of their properties, on the explanation of epistemic and conceptual phenomena.*

RÉSUMÉ. — *Je distingue deux variétés de conscience de soi. J'appelle la première "conscience de soi perspective". Je rends compte de sa nature et j'analyse sa relation aux éléments suivants : le test du miroir de Gallup ; l'immunité à l'erreur d'identification selon Shoemaker ; la possession par le sujet conscient de l'idée d'une pluralité d'esprits ; et quelques-unes des idées de Sartre sur ce que c'est que se concevoir soi-même comme objet. J'appelle "conscience de soi réflexive" une deuxième variété de la conscience de soi. Je rends compte de sa nature et de sa signification épistémologique, et je considère les manières dont la coopération de la conscience de soi perspective et de la conscience de soi réflexive est une condition nécessaire à la réalisation par le sujet pensant de certains objectifs épistémiques. Je conclus par quelques réflexions sur l'importance de la conception métaphysique des sujets de la conscience et de leurs propriétés pour l'explication de phénomènes épistémiques et conceptuels.*

My aim is to characterize two varieties of self-consciousness, which I will call perspectival self-consciousness and reflective self-consciousness. My intention, after characterizing the two varieties, is to explore their psychological and epistemological significance, their relations to one another, and their significance for some wider issues in metaphysics.

It is helpful in this territory to distinguish three kinds of conscious subjects. First, there are what we can call mere subjects, who enjoy conscious states and, as subjects, have some awareness of how they were recently. The content of such awareness arguably involves some form of non-conceptual analogue of the first person. Much animal awareness may be like that. Second, there are subjects of

consciousness who employ the genuine first person concept. The first person concept is something that features in the content of judgements that can be made for reasons. Third, there are subjects of consciousness who are self-conscious, who employ the first person concept in ways that go beyond what is minimally required for use of the first person concept in physical and mental ascriptions. These forms of self-consciousness involve a wider conception of oneself in the physical and mental worlds. My two target varieties of self-consciousness are both subvarieties of the states of this third kind of subject. They are certainly not the only varieties – I think there is much further work to be done in this area. But I do think the two varieties are fundamental to understanding the other, more sophisticated, notions that fall within this third kind.¹

Self-consciousness features as such in our everyday psychological thought, when we appreciate its involvement in such emotions and traits as pride, embarrassment, shame and arrogance. Obama employed the notion in response to the question of whether living in the White House would spoil his young daughters. He replied, 'Right now, they're not self-conscious'.² In ethology, some variety of the notion is employed when it is debated whether passing some form of the mirror test is sufficient evidence for self-consciousness (an issue I will touch on soon). Self-consciousness, and the states it makes possible, are crucial to understanding Romantic thought and sensibilities. In philosophy, some form of the notion is central to the thought of such diverse thinkers as Kant, Hegel, Kierkegaard, Sartre and Strawson – to name but a few. Though I cannot discuss all of the preceding, I do hope to characterize the two target varieties of self-conscious-

1. It is important here to distinguish those issues that are terminological from those that are not. In *The Varieties of Reference* (Oxford: Oxford University Press, 1982), G. EVANS uses 'self-conscious' for the kind of thinking that employs the first person concept, but this, as he says (p. 206), is a mere label, and not a substantive thesis. In fact in the Appendix to his chapter 'Self-Identification', Evans introduces the notion of what he calls 'full self-consciousness', which involves appreciating that in your first person thoughts you are both subject and object of your thought. His discussion of these issues was sadly unfinished, and no final view of full self-consciousness emerges from this chapter and its appendix. But it certainly seems that he was aware that there is at least one notion of self-consciousness that goes beyond plausible minimal conditions for the ascription of the first person concept. In contrast, I suspect my divergence from Sebastian Rödl on some of these matters is more than terminological (despite some other important areas of agreement). In his book *Self-Consciousness* (Cambridge, MA: Harvard University Press, 2007), Chapter 1, Rödl identifies self-consciousness with the capacity for first person thought, and this is not a stipulative matter. I in fact have no strong views about the use of the word 'self-consciousness' in English, or its cognates in other languages. My substantive claim is that however we use these terms, we ought also to distinguish the second and third kinds of consciousness mentioned in the text above, and should take it as a philosophical task to elucidate the subvarieties of the third kind.

2. From *The New York Times*, "First Chores? You Bet" by Rachel L. SWARNS, February 21, 2009: "Those are some special girls, and everyone is rooting for them to make it through this intact," Craig Robinson, Mrs. Obama's brother, said in an interview. The president echoed that sentiment. "Right now, they're not self-conscious. You know, they don't have an attitude," Mr. Obama said on CBS News.

ness in such a way that one can trace some unified common threads through these various areas of thought and feeling.

I. WHAT, IF ANYTHING, IS THE SIGNIFICANCE OF THE MIRROR TEST?

It is natural and illuminating to begin a discussion of perspectival self-consciousness by considering Gordon Gallup's famous mirror test.³ The test involves placing on an animal's forehead a visible chalk mark when the animal is asleep or anaesthetized. In some cases the animal, when presented with itself in a mirror after this marking, is capable of locating the mark on its own forehead with its own limbs, and then wiping it off. Does the presence of this ability establish that the animal has some form of self-consciousness? And if so, what is the correct characterization of the self-consciousness so established?

We need to apply here the distinction between evidential conditions and constitutive conditions. It can, very reasonably, be objected that the ability to use a mirror in removing a mark from one's forehead can be explained without attributing to the animal recognition of itself in the mirror. Grasp of a correlation between what one is doing to oneself and what is seen in the mirror suffices. But that is not the issue I want to pursue here. I want to raise a philosophical question about the case in which the animal does recognize itself in the mirror. So I am focusing on the case in which the animal has a representation with the content "That's me", where the "That" expresses a perceptual-demonstrative mode of presentation individuated by the way an animal is given in the mirror. On what notion of self-consciousness does such self-recognition give a reason for attributing self-consciousness to the animal? How should we characterize that variety of self-consciousness, and does it get a grip beyond the mirror cases?

When a subject learns something about what is on its forehead, or its other bodily features, by looking in a mirror, the subject comes to know something about himself in a way in which he could equally come to know something about some other object or other animal or person. The subject employs a way of coming to know a third person proposition or content. In these examples, that way involves taking at face value the content of his perceptual experience in respect of something thereby perceived. Because this is a third person way of coming to know, it is also a way in which others can come to know about the

3. G. GALLUP, Jr. "Chimpanzees: self-recognition". *Science* 167 (1970): 86-87. For a more extended discussion, see Gallup's *Self-Recognition in Chimpanzees and Man: A Developmental and Comparative Perspective* (New York: Plenum Press, 1979).

subject. But the subject can employ this third personal way of coming to know without having in advance a conception of other subjects of experience who are knowers.

In these mirror cases, the subject comes to know such a proposition as "That animal has a mark on its forehead" and uses an identity "That animal is me" to infer "I have a mark on my forehead."

So we can formulate what we can call the *mirror-motivated criterion* for this variety of self-consciousness. The mirror-motivated criterion states that an exercise of this form of self-consciousness consists in using knowledge of the form "That G is F" (or some other third personal mode of presentation in place of *That G*) and "That G is me" to infer "I am F."

In cases that meet this mirror-motivated criterion, the perception that gives the reason for making the judgement is not a subject-reflexive state in which the subject is represented as (for instance) himself having a mark on his forehead. That is, the perception of something as having a mark on its forehead is not one whose fundamental correctness condition requires, *de jure*, by virtue of the nature of the way the objects are given in perception, the subject who is enjoying the experience to have a mark on his forehead.⁴ That is why acceptance of the identity "That G is me" is needed to reach from the perception knowledge or registration that represents oneself, as such, as having a mark on one's forehead. These are paradigm cases in which the way of coming to know something about oneself is not immune to error through misidentification.⁵

There seems to be a general kind of self-consciousness of which the ability to recognize oneself in a mirror, and thereby gain knowledge of oneself, is a particular manifestation. It is this general kind or variety that I am labeling as perspectival self-consciousness. Is the mirror-motivated criterion a satisfactory way of capturing what is distinctive of this more general notion of perspectival self-consciousness? At a very crude and intuitive level, one wants to say that perspectival self-consciousness involves a certain appreciation. There is a distinctive way in which it can be given that things other than oneself have properties and stand in various relations, and perspectival self-consciousness seems to involve the appreciation that one can oneself have those properties and stand in those relations too, as those properties and relations are given in that distinctive way. A large part of the challenge is to explain what the "distinctive way" is.

The idea of perspectival self-consciousness should not be collapsed into the

4. For further discussion of the notion of subject-reflexive states and events, see section 4 below; and my paper 'Subjects and Consciousness', forthcoming in *Self and Self-Knowledge*, ed. A. Coliva (Oxford: Oxford University Press).

5. S. SHOEMAKER, 'Self-reference and self-awareness', repr. in his collection *Identity, Cause and Mind: Philosophical Essays* (Cambridge: Cambridge University Press, 1984).

simpler formulation that the perspectively self-conscious subject appreciates that he is capable of having the same properties as other things or subjects. That simpler condition is met by someone who is capable of judging that he is spatially close to some object, and who is also capable of judging that other things are spatially close to various other objects. This falls far short of perspectival self-consciousness. The simpler condition is also met by someone who is both capable of judging that he himself is in pain and capable of judging that others are in pain too. Such a subject may be merely capable of first personal thought, and be conscious. But it is not in virtue merely of meeting the simpler condition that the subject is thereby self-conscious. So we really do need to take on the task of specifying the "distinctive way" more clearly.

The major problem with the mirror-motivated criterion is that it treats only a very special case, a case that is not essential to perspectival self-consciousness. We have no difficulty in conceiving of a world in which there are no mirrors and no reflecting surfaces, no depictions, and no videos. In such a world, no subject can truly think an identity "That G is me", where "that G" expresses a perceptual demonstrative of the same general type that could be used in perceiving someone else's face. But it seems that there could be what we would classify as perspectively self-conscious subjects in such a world without visual reflections. So what less restrictive criterion would these subjects meet? It is also plausible that congenitally blind subjects can be perspectively self-conscious. What is the more general condition of which the mirror-motivated criterion is merely a special, and inessential, case?

II. A WIDER CRITERION FOR PERSPECTIVAL SELF-CONSCIOUSNESS

As a step towards formulating a more general condition for perspectival self-consciousness, I start by drawing a distinction. I continue to use the terminology under which a predicative concept is a mode of presentation of a property.

We can draw a distinction, amongst concepts, between those whose understanding-condition makes special reference to the thinker's own possession of the property picked out by the concept, and, by contrast, those concepts whose understanding-condition does not make any such special reference to the thinker's own possession of the property. Anyone who holds that grasp of the concept *is in pain*, or grasp of the concept *has an experience of redness*, involves special reference to the thinker's willingness to apply these concepts to himself when he himself is, respectively, in pain or is enjoying an experience of redness, will place these concepts in the first rather than the second of these categories. In

such cases, the thinker's grasp of the concept, and its applicability in the world, gives a special place to his knowledge of what it is for he himself to fall under the concept (and so to have the property picked out).

In saying that the understanding-condition makes special reference to the first person application of the concept, I offer a formulation that is neutral between various substantive theories of concepts and of concept possession. I myself hold that concepts are individuated by their fundamental reference rules, and that grasp of a concept is tacit knowledge of its fundamental reference rule. Under that approach, the fundamental reference rule for a concept like *pain* will mention the case in which the concept applies to oneself, and relate the correctness of applications to others to that first person case. Under conceptual-role treatments of concepts, the special reference to first person application will be explicit in the formulation of the proposed concept-individuating conceptual role, which will treat first person (present tense) application in a separate clause. As far as I can see, the distinction between understanding conditions that do, and understanding conditions that do not, make special reference to first person applications of the concept will be available under any of the current contenders amongst theories of concepts and understanding.

Unlike the understanding-conditions for the concepts *pain* and *experience of red*, the understanding-condition for an observational shape concept does not make special reference to the first person (and present tense) application of the concept. The understanding condition for an observational concept will indeed make special reference to the thinker's perception of something as having a certain shape; but it will not make reference to *his* having that shape. The same applies to concepts of mass, of temporal intervals and duration, concepts of number. In none of these cases will the understanding-condition make special reference to first person applications of the concept in question. I label concepts whose understanding-condition makes special reference to first person application *concepts anchored in the subject*.

I propose that a necessary condition for perspectival self-consciousness is the capacity to know propositions of the form *I'm ϕ* , for some range of concepts ϕ that are not anchored in the subject. When a subject meets this condition with respect to such a concept, I say the subject is perspectively self-conscious with respect to that concept.

This necessary condition helps to capture the "distinctive way" I mentioned earlier. On this account, a perspectively self-conscious subject is capable of knowing he falls under concepts for which his fundamental understanding is not given in terms of what it is for *him* to fall under them. It is natural to cast the point in terms of the perspective a thinker has to have on an object or property in order to think of it in a particular way. If we are allowed to do that, we can say:

perspectival self-consciousness involves a subject thinking of himself as falling within whatever perspective on an object has to be employed in coming to know that it is ϕ , where grasp of the concept ϕ is not given to him in terms of what it is for he himself to be ϕ .

This is a stronger condition than merely being capable of judging both of himself and of other things that they fall under a given concept. To meet the stronger condition, the subject must conceive of himself as placed in the world in such a way that he himself meets the same kind of condition for a concept ϕ , not anchored in the subject, to apply to himself as can be met by other objects to which the concept ϕ applies. Since the subject has to have a conception of himself that is not given simply by his grasping first person contents, I call this necessary condition for perspectival self-consciousness a contribution towards a *reorientation account*.

The necessary condition concerning concepts not anchored in the subject includes cases covered by the mirror-motivated account, but it also applies more widely.

Let us take the claim of inclusion first. A subject who comes to know 'That's me' by seeing himself in the mirror, and so comes to learn new things about his body, will meet the necessary condition concerning concepts not anchored in the subject. The properties whose application to yourself which you come to know about by looking in the mirror – properties specifying your appearance, the state of your face, your girth, the state of your clothes and of your hair – are all thought about by you under concepts that are not anchored in the subject.

As we required, the necessary condition that is a step towards a reorientation account also applies more widely than the mirror-motivated account, because its requirements can be fulfilled in a world without mirrors and reflecting surfaces. A subject may in such a world still stand next to a small tree, feel that its top is above his own head, and thereby learn *That tree is taller than me*. The subject may alternatively make two marks on a wall level with his own and with the tree's height, and see that the mark for the tree is above his own; and so forth for indefinitely many other ways of coming to learn *That tree is taller than me*. The concept *taller than* is not anchored in the subject – its understanding-condition does not make special reference to its application in the first person case. In such a simple case, by feeling or marking the difference in height between himself and another object, our subject employs a method of determining difference in height that is applicable to objects he can reach or touch. He simply applies the method to something other than his own body. Similar points could be made for such concept as *having a certain kind of face, having a certain profile, having a certain hairstyle*.

Even if someone is unaware of her current properties – as someone may be

unaware of the pattern of her hair at a certain time, or, in a more extreme case, may be unaware of her height, like Alice at one point in Lewis Carroll's story – she may nevertheless be aware that she has some property in a certain range. This is still a variety of perspectival self-consciousness. One can know that one has some property or other within a certain range simply as a consequence of knowing that one is a material object. This unspecific perspectival self-consciousness is still actual, rather than being merely a state of potential self-consciousness. The subject has the capacity mentioned in the definition of perspectival self-consciousness without currently exercising it.

The examples I have given of concepts anchored in the subject have been concepts of psychological properties. Some spatial concepts are also anchored in the subject. The relational concepts *x is to the left of y* and *x is to the right of y* are very plausibly anchored in the subject. A thinker's grasp of them primitively involves understanding of what it is for something to be to the left, or to the right of him. His understanding of other instances is that they involve the same spatial relation as is instantiated in these first person cases.

Is the capacity to know propositions of the form *I'm φ* for some concepts φ not anchored in the subject not merely a necessary, but also a sufficient, condition for perspectival self-consciousness? I do not think it is. Concepts such as *x is a certain distance from y*, *x is in front of y* are not anchored in the subject. Consider knowledgeable first person ascriptions of these made on the basis of perception, in contents such as *That chair is more than a foot from me*, or *I am in front of a house*. These judgements are not, intuitively, exercises of perspectival self-consciousness. They do not, intuitively, involve taking some form of third person perspective on oneself.

Why are these spatial cases not exercises of perspectival self-consciousness, even though the concepts self-ascribed in them are nonetheless not anchored in the subject? I am inclined to think that these are not cases of perspectival self-consciousness because the rationality of making a specifically first person ascription in these cases is adequately explained by the conditions for something to be my body, and the fact that, for example, for *That chair is more than a foot from me* to be true is for it to be the case that *That chair is more than a foot from my body*. The condition for something to be my body is that it is the one from which I perceive, when all is functioning properly, it is the one that moves when I try to move, and so forth.⁶ This condition and the equivalence of the relevant first person proposition with a proposition about my body suffice to explain the rationality of judging the first person content *That chair is more than a foot from me* when I perceive a chair to be more than a foot from my body. (We are not

6. Cp. S. SHOEMAKER, 'Embodiment and Behavior', repr. in his *Identity, Cause and Mind*.

concerned here with explaining the rationality of taking perception at face value, but only with the difference in the rationality of accepting different propositions in the light of perception, *modulo* the general default rationality of taking perception at face value.) When the rationality of making such self-ascriptions can be so explained, the subject is exercising no more than a first person perspective on himself. It is for this very reason that the ability to enjoy primitive spatial perceptions with contents concerning the layout of the subject's immediate environment involves only first person representation and consciousness. It does not in itself involve self-consciousness. Equally, self-applying concepts that are anchored in the subject seems to involve only consciousness, and not any kind of perspectival self-consciousness. These verdicts seem to be in accord with intuition. More generally, given that any use of the first person concept refers to a subject, it should not be surprising that some epistemic features of first person contents trace back to features of the constitutive account of what it is for a subject to have a location and body in a spatial world.

So the full necessary and sufficient conditions for perspectival self-consciousness on the reorientation account that I offer are these: the subject must be capable of coming to know propositions of the form *I'm φ* where φ is not anchored in the subject, and where the rationality of making these specifically first person judgements in the cases in which the subject comes to know them is not fully explained by what it is for something to be the subject's body, and by *my body/I* equivalences. No doubt a subject realized by a brain in a vat could in some sense enjoy perspectival self-conscious too, but this possibility seems parasitic on the conditions just given. For such a subject, it is as if he enjoys perspectival self-consciousness as described.

The notion of capacity in this characterization of perspectival self-consciousness is to be understood relatively broadly. No doubt someone who has real reason to think there is an evil demon around, or real reason to think there are all sorts of misleading evidence, may not be in a position to know the relevant propositions of the form *I'm φ*, even if he is perspectivally self-conscious. But he is in a position to know them if he didn't have these real reasons for doubt, and in that sense has the capacity required in the criterion for perspectival self-consciousness. I should also add that I have formulated the criterion in terms of knowledge because my own view that there is a close nexus between concepts and the conditions for certain kinds of knowledge.⁷ But theorists with different approaches to concepts, or to knowledge, could adapt these ideas to their own framework. One could, for instance, adapt the approach given here to warrant-

7. See my *Being Known* (Oxford: Oxford University Press, 1999), Chapter 2.

based treatments, and characterize perspectival self-consciousness in terms of a rational sensitivity to warrants for the relevant propositions of the form $I'm \phi$.

It is very plausible that judgements of the form $I'm \phi$ that are, under this reorientation account, exercises of perspectival self-consciousness are also vulnerable to Shoemaker-style error through misidentification relative to the first person constituent of their content. If I come to know *I have a Roman nose* by looking at a shadow of a nose, or come to know my height from a mark made on a wall, or come to know *I have a bad haircut* by looking at a photograph, these are clearly cases in which I could come to know, by these same methods, that someone has a Roman nose, or has a certain height, or has a bad haircut, but be mistaken that it is I who have these properties.

Why does perspectival self-consciousness as characterized by the reorientation account imply vulnerability to such errors of identification? I suggest that there are two ways in which immunity to error through misidentification can arise, and in the nature of the case neither of them is present in instances of perspectival self-consciousness as characterized in the reorientation account. In one kind of case, immunity to error through misidentification flows from the nature of the concept that is self-ascribed, that is, in the case in which the concept is anchored in the subject. The very account of what it is for something to fall under such a concept as *is in pain* will, if the subject applies the concept in a way suitably corresponding to the first person clause in the understanding-condition, ensure that it is the subject who falls under the concept.

The other kind of case is that in which the constitutive account of what it is for a predication $I'm \phi$ to be true, for a certain range of concepts ϕ , also ensures that when such a predication is made for certain reasons, the concept applied will be true of the thinker. That is the case illustrated by such spatial predications as *That chair is more than one foot from me*.

Now the very characterization of perspectival self-consciousness ensures that the self-ascriptions it involves are cases of neither of these two kinds. If there is no other source of immunity to error through misidentification, then every self-ascription that is an exercise of perspectival self-consciousness will be vulnerable to error through misidentification. The two kinds of case of immunity to error have their sources in the nature of the predicative concept and the nature of the subject concept involved in a judgement $I'm \phi$.

Why, it may be asked, should we not simplify the criterion for perspectival self-consciousness by formulating it immediately in terms of vulnerability to error through misidentification? Can we not say that a perspectivally self-conscious subject is one capable of coming to know contents of the form $I'm \phi$ in ways which are vulnerable to error through misidentification?

My principal reason for rejecting this suggestion concerns the intended expla-

natory status of the account I have been presenting. The account in terms of concepts anchored in the subject and judgements made reasonable by the account of what it is for a subject to have a location and to have a particular body is plausibly explanatory of some of the boundaries of immunity to error through misidentification. The account traces the phenomena to those sources. There are other cases of immunity to error through misidentification, outside the first person examples. *This book is blue* is immune to error through misidentification, relative to the constituent *This book*, when it is judged in ordinary circumstances on the basis of a perceptual experience of the book as blue, the same experience which makes available the perceptual demonstrative *This book*. These cases of immunity to error through misidentification have their explanation too, one which traces back to what makes something the reference of the perceptual demonstrative *This book* and the nature of the observational concept *is blue*. In offering an explanation of the boundaries of immunity to error through misidentification in the first person cases in terms of features of the first person concept and of the predicative concepts in question, we offer an explanation of the same structural character as applies in the explanation of such immunity in the cases beyond the first person.

I close this section with two comments relevant to assessing the role and significance of perspectival self-consciousness. First, prima facie perspectival self-consciousness can be present in subjects who neither exercise nor possess concepts of mental states and events, whether their own or others. The exercises of perspectival self-consciousness in the examples of coming to know *I have a Roman nose*, or a certain facial profile, or a certain height from markings on a wall do not require me to have concepts of perceptual experience. They require only a certain rational sensitivity to my experiences in coming to make judgements about the non-mental world.

The second comment concerns the fact that the examples of perspectival self-consciousness I have been using involve spatial and material properties and a subject's self-ascription thereof. One of our intuitive notions of self-consciousness runs far beyond these cases. We speak in everyday discourse of someone being self-conscious in respect of how he is thought about by another particular person, or by his friends, or by his colleagues, or in respect of how he is represented in the academic literature, or in the press, or on the web. As far as I can see, these are all phenomena that can be accommodated by the reorientation account, once we recognize that the concept ϕ involved in the self-ascription may be *thought of as irritable by his friends*, or *represented as competent in the press*. These concepts are not anchored in the subject. Knowledgeable self-ascriptions of them can be exercises of perspectival self-consciousness in the sense of the reorientation account. There is correspon-

dingly a generalized notion of perspective in each of these more specialized varieties of perspectival self-consciousness: the perspective given by the view of the subject's friends, or of the representations of the press, and so forth.

III. THE REORIENTATION ACCOUNT AND RELATIONS TO OTHER MINDS

Perspectival self-consciousness with respect to a range of properties prepares the way for a subject's ability to think of himself as an object of others' awareness. It prepares the way, but it does not suffice for that ability. Perspectival self-consciousness is, on the characterization I have given, something more primitive than having a conception of oneself as the object of others' consciousnesses, because perspectival self-consciousness as defined does not involve the thinker's having the conception of many minds, of a range of other subjects of consciousness, at all. The methods you employ in coming to know *I'm ϕ* in an exercise of perspectival self-consciousness, where ϕ is a spatial or material concept, do not (to all appearances) involve or presuppose that you have a conception of other minds.

When, however, a thinker does have a conception of many minds, and is perspectivally self-conscious with respect to a concept, these two can be combined to make available a conception of herself as the object of others' perception that she falls under the concept. When you come to know the arrangement of your hair by looking in a mirror, or learn of your comparative height from the marks on the wall, you are in a position to appreciate that the methods which you employ could, suitably adapted, be used by other people to come to know that you have these properties.

There are some claims in Sartre's writings that seem to contradict what I have just said. Sartre writes, for example, in a passage in which "the Other" means another subject of consciousness: "...for how could I be an object if not for a subject? Thus for me the Other is first the being for whom I am an object; that is, the being through whom I gain my objectness. If I am to be able to conceive of even one of my properties in the objective mode then the Other is already given".⁸ Sartre's views in the general area of subjecthood and consciousness of oneself as a subject are of great interest, and deserve more attention than I can give them here. But Sartre's claim in the passage just quoted, if taken at face

8. *Being and Nothingness*, Part III 'Being-for-others', Chapter One 'The Existence of Others', Section IV, 'The Look', at p. 270 in the translation by H. Barnes (New York: Philosophical Library, 1956); at p. 317 in the original French edition *L'Être et le Néant* (Paris: Gallimard, 1943).

value, seems too strong. If we are concerned with a sense in which the only way I can be an object is to be an object for a subject, that sense presumably concerns being represented, either in consciousness or thought, as an object. Under that reading, Sartre's rhetorical question – "How could I be an object if not for a subject?" – invites a straightforward answer: in perspectival self-consciousness, there is indeed a subject for which one is (represented as being) an object; but that subject is oneself.

Sartre could reply that in perspectival self-consciousness as I have characterized it, the thinker is merely thinking of himself as object, as falling under concepts not anchored in the subject; the thinker is not perceptually conscious of something both as himself and as object. Sartre's insistence on the distinction between what is thought and what is in perceptual consciousness is one of the best features of his writing, and it needs to be respected. Nonetheless, I do not think the reply just drafted for Sartre would be a defence for him. Consider Sartre's famous example of the person bending down to spy through a keyhole, who suddenly realizes that he is being seen by another person in this compromising posture and purpose.⁹ Though Sartre considers the example to be one in which, as he would say, I am an object, viz. one in which I am aware of someone else perceiving me, it is not really true in this example that I have a perceptual awareness of myself as object. I have, rather, a perceptual awareness of someone else perceiving me as an object (as well as perceiving me as being a subject).

It is indeed true that it is impossible for something to be given in consciousness in such a way that it is thereby both a presentation of something as an object – by which I mean, presented in such a way that it's potentially informative that it's me – and simultaneously a presentation of something as me. This is arguably an insight that is common to Hume, Kant, Fichte and Wittgenstein – another worthy topic for discussion on some other occasion. But that agreed impossibility cannot be used in support of the present aspect of Sartre's position, for the introduction of the Other into the phenomenology does not make something possible which would otherwise be impossible. It is still impossible after the introduction of the Other. The introduction of the Other gives us a case of a subject's conscious awareness of another perceiving him as an object; that remains distinct from perceiving himself, given as himself, also as an object.

The full realization by the subject looking through a keyhole that he is being seen by another person is available only to someone who is capable of thinking of his posture in a way that is not completely anchored in the subject. He has an observational way of thinking of a person bent down and looking through a

9. *Being and Nothingness*, p. 259ff; p. 298 in the French original.

keyhole. That observational concept is not anchored in the subject (which is not to say that his own knowledge that he falls under it is observational). This observational concept of this particular posture is constitutively prior to the conception of other subjects who may also apply this spatial concept and may perceive it to be instantiated. In the embarrassing case, another subject perceives it to be instantiated by oneself. Such spatial concepts make possible detailed and specific psychological predications of other subjects in respect of what they are seeing, when a thinker has the conception of other perceiving subjects. That is the order of philosophical and constitutive explanation, rather than the converse.

Other theses may be at stake in the passages from *Being and Nothingness* that I have been considering. At least two other readings are possible (and perhaps Sartre intended all three). A second possible construal is that the passages in question propound the thesis that what may seem to be first person representation is not genuinely referential, referring to a subject, until the Other is recognized. Sartre says that what I get from the Other is "the abstract moment when the self is apprehended as an object"; and that "I must obtain from the Other the recognition of my being".¹⁰ I do not reject this as a construal, but I do dispute the truth of the thesis. I would argue that for fundamental metaphysical reasons, there is no making sense of attribution of mental events and states except as possessed by a subject. Further, even in the case of some of those states that Sartre would describe as "non-thetic consciousness", where he agrees there is consciousness of one's own consciousness, but no reference to a subject, I would say that these are subject-reflexive events and states in the sense mentioned earlier. They have a rather primitive kind of content made possible by the nature of subjects themselves. Nevertheless they still refer to the subject who has the consciousness in question. I do not think we have any account of the correctness conditions of the contents of such consciousnesses that does not involve reference; and the reference must be to subjects.¹¹

A third possible construal of these passages is that the subject has no conception of himself *as a subject* until he has the capacity to perceive the Other as perceiving him. This is a thesis worthy of consideration in its own right. (It should be added, though, that first person thought is enough for the Cogito to get off the ground. So either Sartre should not have associated the distinction he has in mind with the possibility of employing the Cogito; or else this third construal fails as interpretation.) The first question for this third construal is: what does "*as subject*" mean in its formulation?

10. *Being and Nothingness*, pp. 236-7; pp. 281-2 in the French edition.

11. For additional considerations, see the discussion of Anscombe's views in my *Truly Understood* (Oxford: Oxford University Press, 2008), Chapter 3; and my 'Subjects and Consciousness'.

One answer would be that it means that the subject is employing a one-place concept ζ *is a subject*; the subject applies it to himself; and the substantive thesis is that a perception of the Other is what makes this possible. The problem with accepting this as a true thesis (as opposed to correct interpretation of Sartre) is that in thinking of someone as the Other, one must already be thinking of him as a subject. For this, one must have some conception of being a subject. It is highly plausible – and indeed, a principle accepted by Sartre himself earlier in *Being and Nothingness* – that other subjects are conceived of as things of the same very general kind as oneself.¹² But that account of being a subject makes it immediately obvious, and (for those who employ the notion) a priori, that I am a subject. My appreciation of that truth does not, on this account of grasp of the notion of subjecthood, need to proceed via perception of the Other.

Another answer to this question about the third interpretation is that thinking of oneself "*as subject*" means thinking of oneself as someone who is not merely employing the first person in thought, but thinking of oneself as employing the first person. This interpretation is invoking the entirely appropriate distinction between using the first person way of thinking, and thinking about that way of thinking, as the first person way of thinking.¹³ As a thesis, it is open to the objection that various forms of reflective self-consciousness, that involve thinking about one's own conscious mental states and events in ways made available by their being one's own conscious mental states and events, already involve attributing to oneself use of the first person way of thinking, as the first person way. It is not obvious that the capacity for this self-attribution involves perception of the Other perceiving oneself. I defer to a later part of this paper further consideration of reflective self-consciousness.

I return now to perspectival self-consciousness in its own right. If indeed other subjects are conceived of as things of the same kind as oneself, then they will stand in the same sorts of relations to things, and to properties and relations instantiated in their environment, as the subject himself stands to things, and to properties and relations, instantiated in his own environment. Perceptual relations constitute an important subclass of such relations. So when a thinker comes to know of a concept, not anchored in the subject, that she falls under

12. "The Other is a thinking substance of the same essence as I am, a substance which will not disappear into primary and secondary qualities, and whose essential structure I find in myself": *Being and Nothingness*, p. 223, at the start of the section entitled 'The Reef of Solipsism'; p. 267 in the French edition. This seems to me to be the same fundamentally first personal account of grasp of the general concept of subjecthood as I proposed in *Truly Understood*.

13. The distinction is crucial both to drawing significant distinctions in the philosophy of mind, and to an adequate theory of indexical and demonstrative modes of presentation. See my paper "Demonstrative Thought and Psychological Explanation", *Synthese* (Symposium on Indexical and Demonstrative Reference) 49 (1981), 187-217.

that concept, she will be employing a kind of method that will be available to other subjects of consciousness in coming to know about her. She will also be in a position to appreciate that others can use that method to come to know about her. Again, this is the direction of explanation, rather than that asserted by Sartre in his treatment of his example.

Perspectival self-consciousness is thus a necessary condition for a wide range of familiar psychological phenomena that we describe in everyday discourse as involving some intuitive notion of self-consciousness. The psychological states of embarrassment, shame, and certain kinds of arrogance all involve the conception of others' awareness of oneself as having certain properties, where these properties are thought of, by the subject of these states, under concepts that are not anchored in the subject.¹⁴ For those interested in a particular form of Romantic fantasy, as exemplified in one of Heinrich Heine's poems in his *Buch der Lieder*: to form the idea of a Doppelgänger of oneself requires one to have perspectival self-consciousness with respect to some of one's properties. We can give the following continental-style formulation. One can be aware that one is an object of a certain kind for others only if: one has an awareness of that kind that involves an appreciation of its instantiation by oneself as given in a perspective that is not purely first personal.

This combination of perspectival self-consciousness with a conception of oneself as an object of other's consciousness is a crucial element in articulating what is right in Kierkegaard's treatment of some of his discussions of the self as involving a certain kind of self-consciousness. There are insights here that are completely independent of Kierkegaard's religious inclinations. In *The Sickness unto Death*, he writes: «this self takes on a new quality and qualification by being a self directly before God (p. 79);¹⁵ [...] what an infinite accent falls on the self by having God as the criterion! The criterion for the self is always: that before which it is a self, but this in turn is the definition of "criterion".» (p. 79)

Kierkegaard's conception of the self as being 'before' some other subject, another person or God, who evaluates it, involves the idea of a person, and his properties, being known about by another. One must feature in the other's, the evaluator's, perspective. For a subject to have this conception of herself as so evaluated, she must have the conception of at least some of her properties as available from the other's point of view. She must conceive of some of her properties in such a way that there are ways of knowing of them that are not only

14. For further discussion of these states, see L. O'BRIEN'S 'The Ordinary Concept of Self-Consciousness', forthcoming in *Consciousness and the self*, ed. J. Perry and J.L. Liu (Cambridge University Press, forthcoming).

15. Translation by E. Hong and H. Hong (Princeton: Princeton University Press, 1980); and similarly for later quotes from this work.

available to she herself. Perspectival self-consciousness in respect of the relevant range of properties makes available this conception.

We should also draw a conceptual distinction between the case in which the other perceives one as having certain properties that are not psychological – one's spatial and material properties, for instance – and that in which the other perceives one as having properties that only a subject can have – such as acting and perceiving. Kierkegaard's discussion, and also our ordinary states of embarrassment and pride, both involve the second of these as well as the first. (So does the conception of a genuine Doppelgänger. Since one is oneself a subject, so must one's double be so too.) But even awareness that someone else knows some of your material and spatial properties already involves your having some perspectival self-consciousness in respect of some of your properties and relations.

What is the relation between what have become known as the mirror-neuron phenomena and perspectival self-consciousness? Humans have the ability to think of a bodily action in a way that has special ties to both action and perception. A subject may perceive a particular gesture in a certain way on a particular occasion. Having so perceived it, our subject does not need to engage in any conscious and personal level inference to act in that same way, to imitate it. Conversely, if a human subject makes a gesture on a particular occasion, he needs no conscious and personal level inference to recognize that an action of another human is of the same kind as he has just made.¹⁶ We can call these special ways of representing actions "bouncing ways", since they permit perceived and produced actions of given type represented in one of these ways to bounce back and forth without the need for conscious inference between the contents of action and the contents of perception. We can now pose the question: is it necessarily the case that anyone who is capable of employing these bouncing ways is also perspectively self-conscious?

I answer this question in the affirmative, and offer this argument in support. Suppose a subject sees another's bodily movement as tracing out a certain shape in space. Perceiving something as tracing out, or as occupying, this shape does not in itself involve the ability to represent shape in one of the bouncing ways. This aspect of the perception just involves spatial perception – although of course other features of the percept may involve much more than spatial perception. The concept of tracing out or occupying that particular shape, whether the shape is thought of demonstratively ("that shape", made available by the recent perception), or as recognitionally (V-shaped, oval-shaped, etc.), is not a concept

16. For an engaging introductory overview, see M. IACOBONI, *Mirroring People: The New Science of How We Connect with Others* (New York: Farrar, Strauss and Giroux, 2008), Chapter 1.

anchored in the subject. An account of what individuates these shape concepts, whether demonstrative or recognitional, has no special clause treating the case in which the thinker herself is V-shaped, or oval-shaped, or has that (demonstratively given) shape, or is making that shape in some action. But when a subject is capable of employing these bouncing ways of representing actions, he will have a way of knowing a content of the form *I am making a gesture with a path of such-and-such shape*, where the gesture is thought of both under the spatial way of thinking, not anchored in the subject, and in the bouncing way. This means he has a way of coming to know that he is making a gesture tracing a certain kind of spatial path, where this last is, as we said, not anchored in the subject. So he meets the reorientation criterion for perspectival self-consciousness.

Consequently, as we also noted, this subject has a conception of his action-type which, when he has the conception of other subjects, makes available to him the thought that others will be able to perceive it as of a certain type (given under the bouncing way). I am therefore in agreement with those who see intrinsic connections between one type of self-consciousness and the mirror-neuron phenomena. No doubt there are several views falling under that general description. The particular form of connection that I am endorsing is the thesis that subjects who employ the bouncing ways of representing actions will enjoy perspectival self-consciousness with respect to their own actions represented in these ways.

In this respect at least, I am in partial-agreement with V. S. Ramachandran's remark that the mirror-neuron phenomena may be "the dawn of self-awareness".¹⁷ This may not be the only route to perspectival self-consciousness, and it is certainly not in principle the only route. But it may be an actually existing, psychologically real route humans employ.

Perhaps surprisingly, not all social and psychological interactions a subject may have with other subjects, not even relatively simple ones, require the subject to exercise the capacity involved in perspectival self-consciousness. This point applies to some of the original cases of joint attention. We can consider two kinds of case.

Take first the case in which you – we can imagine you are in the position of a child – succeed in attending to an object because your caregiver is looking at it. This requires you to look where your caregiver is looking. You can see whether

17. Quoted in John COLAPINTO, Profile of V. S. Ramachandran "Brain Games", *New Yorker*, May 11, 2009, 76-87, at p. 87: "So I made the suggestion that at some point in evolution this system [the mirror-neuron system – CP] turned back and allowed you to create an allocentric view of yourself. This is, I claim the dawn of self-awareness." We also need to distinguish constitutive claims about the connection between allocentric views and perspectival self-consciousness from etiological claims of how such allocentric representations came to exist. I am expressing agreement here with the constitutive element of Ramachandran's view.

your caregiver is looking at you, or elsewhere, or at some particular object, by looking at your caregiver's eyes and head. None of this involves your applying to yourself some concept not anchored in the subject, in the way you would have to if you were required to think of the way in which you feature in your caregiver's perspective on the world during this interaction.

The case does not even need to involve self-ascription on your part of an experience or perception to yourself. You do need to think, of your caregiver "She's looking at *that*". The *that* here involves your use of a perceptual-demonstrative way in which an object or event is given. It does not involve ascribing an experience to yourself. (If the contents here are at a nonconceptual level, there may even be the possibility of having the ability to enjoy representations that attribute seeing a particular object to your caregiver without having the ability to attribute, non-conceptually, seeings to yourself.) But in any case, if you do have a self-ascriptive capacity for perceptions, you do not need to exercise it in successfully engaging in this first kind of joint attention.¹⁸

A second sort of joint attention occurs when you want the caregiver to come to perceive and attend to a particular object, one that you currently perceive, and are thereby thinking of demonstratively as *that so-and-so*. If you are able to distinguish between the case in which your caregiver is attending to that object and that in which she is not, and are able to bring about that the former case occurs, then you will have brought about an instance of this second kind of joint attention. You may need, in the course of bringing this about, to be able to form representations with such contents as *She's not yet looking there, at that; she needs to look further in that direction, at that box*. None of this, either, seems to involve an exercise of the capacity involved in perspectival self-consciousness.¹⁹

The social interactions that do draw on a subject's capacity for perspectival self-consciousness are those that in one way or another involve the subject's being aware that he falls under some concept that is not anchored in the subject. This may be important because the subject needs to think of her partner in the interaction as appreciating that she falls under that concept. Or it may merely be that the subject's own appreciation of her possession of the property picked out by the concept affects how she conducts the interaction. In the former case, the

18. Contrast this with some of J. Bermúdez's descriptions of these cases in his *The Paradox of Self-Consciousness* (Cambridge, MA: MIT Press, 1998), pp. 258-64. This is a minor disagreement compared with a large measure of agreement with the major theses of Bermúdez's book, in particular his insistence on the existence of nonconceptual forms of first person content.

19. Nor indeed does it seem to involve the capacity to self-ascribe perceptions, as opposed to employing perceptually-based representations of objects. Contrast Bermúdez p. 258: "The explanatory requirement to assume that the infant is aware of himself as a perceiver is even clearer in the second form of joint visual attention."

possession of bouncing ways of representing actions can of course be a massive facilitator of successful interaction, particularly in the communication of moods and attitudes by facial expressions.

IV. REFLECTIVE SELF-CONSCIOUSNESS

I now turn to the second variety of self-consciousness, reflective self-consciousness. Though very different from perspectival self-consciousness, reflective self-consciousness still falls under the general umbrella characterization of being a state that goes beyond the minimal capacity for first person thought, even in realm of psychological self-ascriptions.

A first step in characterizing reflective self-consciousness is to characterize more explicitly a notion I touched on earlier, of a state or event being subject-reflexive. A subject-reflexive state or event is one whose content, *de jure* and primitively, refers to the subject who experiences or enjoys that particular state or event. The following are all subject-reflexive in this sense: having a perception in which there is a door to the left of you; enjoying an action-awareness that you are opening the door; remembering being in Athens. These all require, *de jure* and without use of any descriptive material, that for the content of the state or event to be correct, the subject who is now enjoying the state have a certain property: that of having a door to his left, of opening the door, of having been in Athens. This notion of a subject-reflexive state or event is essentially a generic notion of a *de se* state or event, applicable both to some of the states and events of subjects with only minimal first person consciousness, such as knowing one is in front of a desk, and to those with more sophisticated states of self-consciousness. (In the other direction, downwards, my own view is that this notion can also apply both to events with conceptual content and to events with nonconceptual content.)

We can then say: reflective self-consciousness is subject-reflexive awareness of being in a subject-reflexive state, or enjoying a subject-reflexive event, where this first subject-reflexive awareness is based on the second subject-reflexive awareness. Your ordinary awareness that you yourself are perceiving something as to your left, or have an awareness of opening the door, or remember being in Athens, are exercises of reflective self-consciousness in this sense.

This characterization of reflective self-consciousness is meant to identify a significant notion that falls under the characterization of being a state that involves awareness of oneself, as oneself, in its content. It is not designed to be historically faithful to the usage of all those who have, in one way or another,

used the notion of reflective self-consciousness. Indeed, it diverges in extension from some prominent historical usages.²⁰

The criterion is not to be given the reading that the reflectively self-conscious subject has himself to possess the concept of subject-reflexivity. Subject-reflexivity is a modestly theoretical concept that we as theorists use to pick out a certain class of states and events. The subject just represents the state or event as being of the general type it is (a seeing, and so forth, with a certain specific content). The subject thinks of the contents themselves in certain canonical ways.

Simply being in pain or feeling elated are not exercises of reflective self-consciousness by this test, for they do not themselves involve any subject-reflexive state or event, let alone awareness on the part of the subject that he is in such a state or event. A subject-reflexive state has to have a content, and thus a correctness condition, concerning the subject. Being in pain is not something assessable as correct or incorrect. A subject with some representation of ownership will experience certain pains as his own. But neither of these points involves pains being assessable as true or false. In this respect pains differ from perceptual experiences. Nor is perceiving that there's a door in front of one an exercise of reflective self-consciousness. Someone who enjoys such a perception must indeed be capable of some form of first person representation, since the content is of the form *that's a door in front of me*. But perceiving the spatial world that way is not the same as being aware that one is in a subject-reflexive state or enjoying a subject-reflexive event. Subject-reflexive states and events involve awareness of mental states and events. Taking the contents of perceptual experience at face value, even contents with a first person constituent, is not by itself an exercise of reflective self-consciousness.

Reflective self-consciousness requires that the subject-reflexive awareness of being in a certain subject-reflexive state be attained in a certain way. Take a case of reflective self-consciousness involving awareness that one is enjoying a visual experience of a certain kind concerning oneself. This awareness must be rationally produced by the visual experience itself, operating as the subject's reason for self-applying the notion of experience (more particularly, of experience with a given content). We can conceive of a future in which someone can come to know, circuitously and no doubt redundantly, that he is having an experience of

20. Reflective self-consciousness as understood by Sartre in *La Transcendance de l'ego* and in his 1947 lecture *Conscience de soi et connaissance de soi* (both reprinted in the collection under the former title by V. de Coorebyter (Paris: Vrin, 2003)) would include awareness that one is in pain, in which the awareness is of a state of consciousness that does not have the subject-reflexive character required by the present definition. We might call the wider notion reflective subject-consciousness, as opposed to reflective self-consciousness.

a given kind by: seeing a live scan of a functioning brain; knowing the brain is his own; and having psychophysical knowledge that scans of a certain sort underlie experiences with a certain content. Coming to know inferentially in this way that one has an experience of a certain sort would be employing a way of coming to know that would equally be a means of coming to know of someone else that they have an experience of a certain kind. The example would not be a case of reflective self-consciousness in the sense intended.

When a subject judges that he is in pain, in rational response to his own conscious pain, that is a rational responsiveness that has significant features in common with cases of reflective self-consciousness as I have been characterizing it. The rational responsiveness makes appropriate the image of reflecting a state or event, one which exists independently of its being reflected in higher-level awareness or thought. But such rational responsiveness to one's own pain is not itself a case of reflective self-consciousness, because in judging you are in pain, you do not meet the condition of representing yourself as employing the first person. You merely use the first person. By contrast, in the self-reflective cases in which you are aware that you see that you have a door on your left, or in which you are aware that you have an action-awareness of your opening the door, and so forth, you represent yourself as being in a state, or enjoying an event, with a first person content. You do not merely employ the first person notion; you refer to it, and you refer to it as the first person notion. It is perhaps the most primitive form of thinking of yourself as a subject, as opposed merely to being a subject and exploiting that fact in your thought and representations.

V. SOME METAPHYSICS AND EPISTEMOLOGY

The explanation of a thinker's entitlement to make a rational self-ascription of, say, a perception on the basis of the perceptual event itself is in part that the perceptual event is the thinker's own. This last fact about ownership is at the level of the metaphysics of perception and subjects, rather than anything to do with the representation, whether conceptual or nonconceptual, of perception. A perceptual experience enjoyed by the subject himself makes rational the self-ascription of an experience of a certain type, and when all goes well, the occurrence of the perception leads not just to judgement but to knowledge. The transition is of course not inference from a premise about perception. The perception itself is the subject's reason for the self-ascription.

This approach ensures Shoemakerian immunity to error through misidentification of ascriptions so made. The rational transition is from the conscious event itself, and not from some premise identifying someone or other as the owner of

the conscious event. The only case in which such a self-ascription can be made on the basis of a conscious mental event is that in which the conscious mental event is the subject's own. So not only does the self-ascription not rest upon any identity "I am *a*", for some *a* which is antecedently known to enjoy the conscious event. Self-ascriptions made in this way also cannot be mistaken in respect of who it is that is known to enjoy the conscious event.

This simple approach to the relation between the metaphysics and the epistemology of these self-ascriptions contrasts sharply with three others.

First, it contrasts with treatments that aim to offer a philosophical elucidation of ownership of an experience or other conscious state or event in terms of the possibility of some kind of self-ascription of that experience, state or event.²¹ Under those rival approaches, it cannot be a legitimate explanation of the rationality and potentially knowledge-yielding nature of the transition from a conscious state to its self-ascription to say that in the nature of the case, the procedure applies only to the subject's own conscious states and events. That cannot be explanatory on the rival view, because to say that a state or event is the subject's own is simply to say that the subject can, in some specified way, self-ascribe it. What makes the self-ascription rational and potentially knowledge on the rival view would perhaps have to be a matter of contextual self-verification. Suitable self-verifying self-ascriptions, when not based on empirical information, are plausibly cases of knowledge on any reasonable epistemology. So the rival views may have a way of answering the epistemological question of why these self-ascriptions are knowledge. But the account of ownership embraced by these rival views seems wrong. A subject can have experiences, but not be capable of self-ascribing them, simply because the subject does not have any notion of experiences.

Second, the simple model I have offered contrasts in several respects with perceptual models of the knowledge attained in reflective self-consciousness. One obvious respect of contrast, that I have discussed elsewhere, and will not pursue here, lies in the fact that one does not have a further perception of one's perceptions, an experience of them that is distinct from the first-level perceptions themselves.²²

21. There are strong hints of this position in P. STRAWSON, *The Bounds of Sense* (London: Methuen, 1966), especially in the section 'Unity and Objectivity', pp. 97-112. I think the argument of this section would not go through without commitment to the position, even though Strawson's own formulations of the position are somewhat qualified. For further discussion of the position, see J. CAMPBELL, Critical Notice of my book *Sense and Content* (Oxford: Oxford University Press, 1983), *Philosophical Quarterly* 36 (1986), 278-291.

22. See Chapter Six of *Truly Understood*. One actual holder of the perceptual model of reflective consciousness is David ARMSTRONG: see his 'Consciousness and Causality', in *Consciousness and Causality* by D. ARMSTRONG and N. MALCOLM (Oxford: Blackwell, 1984). Kant's talk of inner sense certainly tempts one to attribute the view to him. There are certainly different versions of the model, as Sydney Shoemaker emphasizes in his First Royce Lecture, in his collection *The First Person Perspective and Other Essays* (Cambridge: Cambridge University Press, 1996) at p. 203ff.

Such points against perceptual models of reflective self-consciousness do not, incidentally, rely on anything as strong as the idea, found in Shoemaker, that mental states and events have no nature independent of our ability to know about them in certain ways.²³ The accessibility of first-order conscious mental events and states is, on the present treatment, a consequence of their nature. It is not something to be written into an account of the original nature, as a constitutive condition.

The third, more extended, contrast I wish to draw is with the position Sartre developed in his 1937 work *The Transcendence of the Ego* (at first an article, later a book).²⁴ Sartre's position is summarized in the following quotations:

(a) "the Ego is an object apprehended but also an object *constituted* by reflective knowledge" (p. 34);

(b) "this pole [the Ego – CP] appears only in the world of reflection" (21);

(c) "unreflected consciousness must be considered as autonomous" (19); "It is thus... on this level [of reflection – CP] that egotistic life is placed, and on the unreflected level that is placed impersonal life" (20);

(d) "the Ego is an object that appears only to reflection, and which thereby is radically cut off from the World" (36).

Sartre gives various arguments for his position. Here is a selection (I cannot claim completeness):

(e) "In fact, the Ego never appears except when we are not looking at it. The reflective gaze has to fix itself on the *Erlebnis*, insofar as it emanates from the state. Then, behind the state, at the horizon, the Ego appears. So it is never seen except 'out of the corner of one's eye'. The moment I turn my gaze on it and wish to reach it without going via the *Erlebnis* and the state, it vanishes. The reason is this: in seeking to grasp the Ego for itself and as the direct object of my consciousness, I fall back on to the unreflected level, and the Ego disappears with the reflective act" (39-40).

(f) Sartre also has a thesis about constitution: "what is *really* first is consciousnesses, through which are constituted states, then, through these, the Ego" (34).

23. See his Second Royce Lecture, reprinted in *The First Person Perspective and Other Essays*, pp. 224-5 and p. 240ff. As against Shoemaker's position, I would say about the model of reflective self-consciousness that I have offered here the same as I said in Chapter 7 of *Truly Understood*. There is an explanation of the impossibility of self-blindness (p. 226 in Shoemaker) that does not involve denying the Independence Condition and does not postulate redundant mechanisms of inner sense (see his p. 240). Knowledge of one's own perceptions does not involve perception of those perceptions, but involves rather a rational sensitivity to the occurrence of those conscious perceptual events that is entirely consistent with their having a nature that makes possible and explains, rather than constitutively involves, awareness of those perceptual events and states.

24. Page references are to the translation by Andrew Brown (London and New York: Routledge, 2004), with an introduction by Sarah Richmond.

(g) "It is useless, for instance, if we consider a melody, to suppose there is some *X* which acts as a support for the different notes. The unity stems in this case from the absolute indissolubility of elements which cannot be conceived of as separate, except by abstraction. The subject of the predicate will here be the concrete totality, and the predicate will be a quality abstractly separated from the totality and gaining its full meaning only when it is linked back to the totality.

For these very reasons, I refuse to see in the Ego a sort of *X* pole acting as the support for psychical phenomena. [...] The Ego is nothing other than the concrete totality of states and actions that it unifies..." (29-30).

(h) Sartre is well aware that there are apparently many true propositions involving the first person on the unreflected level. He holds that this appearance can be explained away consistently with his view that the ego does not appear at the unreflective level. His explanation involves the idea of "Body as illusory fulfillment of the *I*-concept" (table, p. 41).

Despite my evident commitment to disagreeing with his principal theses, it seems to me that there are some real insights in Sartre's points. The quotation (e) in particular about the impossibility of finding the ego on the unreflective level is naturally read as putting Hume's famous point about his inability to find himself into the language of a French writer in the 1930s. But as with Hume, I would make the same point again: what is right in Sartre's point is that the ego, as he would put it, cannot be an object of original (non-derivative) attention.²⁵ It does not follow that the ego or subject does not exist; it does not follow that it does not contribute to the individuation of particular mental events; and it does not follow that it has no role to play in an initial characterization of what it is for an event or state to be conscious. From the absence of a certain kind of phenomenology, we cannot soundly draw any conclusions about the ontology of subjects.

Sartre's positive view of the unreflective level sounds like a moderate reductionist view of the ego or subject, of the sort endorsed by Derek Parfit.²⁶ Suppose we accept, for reasons of the metaphysics of consciousness and subjects, that when (for instance) there is an event of perception, there must be an un-reduced subject who is perceiving. Then if reflective thought about the perception is to attribute it correctly, it must attribute it to the same subject as exists already, *pace* Sartre, on the unreflective level. If such metaphysical theses are correct, there is no room for saying that the subject or ego appears only on the reflective level. In fact, on the positive simple account I have been giving, the ownership by the subject, at the unreflective level, plays an important part in explaining the

25. See again my paper 'Subjects and Consciousness'.

26. In his *Reasons and Persons* (Oxford: Oxford University Press, 1986). For discussion of his moderate reductionism, see 'Subjects and Consciousness'.

entitlement, on the part of a subject of a particular experience, to self-ascribe an experience of its relevant kind.

From the standpoint of the metaphysical arguments for an ontology of subjects, Sartre's parallel with melodies, which need no support for the notes that compose the melody, is unconvincing. Any given subject of experience could have had different experiences and other conscious states and events over time. Even the subject's earliest experiences could have been different. A melody, considered as a type, could not go differently. A particular token event of playing the melody could have started and continued differently. But we can make no sense of the token event being a starting of that melody yet also being different in type right from the start. Our notion of the subject of a series of conscious events just is not that of a sequence or sum of conscious events starting with some initial event that is essential to the sequence or sum. It really is, rather, of the subject who has them; and for that particular subject, no particular sequence or initial segment of conscious events or states is required for its existence.

What of Sartre's claim that all uses of the first person in characterizing the unreflective level can be replaced by reference to the body? One can see what he means: at least in the case in which I am embodied, for it to be true that I am in London is for my body to be in London. But this creaks at many points as an attempt at elimination of the first person in characterizing the unreflective level. The first person is still there in the designation "my body". It cannot be eliminated if we can make sense of a subject switching bodies. The proposal does not cover cases in which I have a perceptual point of view in the world but no body with which I can act or feel. The proposal does not do well with ascriptions of action. "I am raising my arm" does not mean "This body is raising this arm". (Under the required subject-free reading, the latter would be true when the arm's rising is a reflex, while the former would be false.) "I am asserting that *p*" does not mean "This body is asserting that *p*"; and so forth. So I disagree with Sartre's claim that "the body and bodily images can consummate the total degradation of the concrete *I* of reflection to the "*I*-concept" by functioning for the *I*-concept as its illusory fulfillment" (90) (though it would be nice to formulate the issue so stylishly).

VI. RELATIONS BETWEEN REFLECTIVE SELF-CONSCIOUSNESS AND PERSPECTIVAL SELF-CONSCIOUSNESS

Perspectival self-consciousness and reflective self-consciousness can be seen as meeting the different needs of two sorts of inquiry a thinker may make about the basic case in which he comes to accept a content *I'm φ*, where *φ* is a concept

anchored in the subject. One sort of inquiry the thinker may make about himself is: what other properties and relations do I have besides *φ*? What sort of thing in the world am I? Answering these questions, questions about the reference of the first person, requires the thinker to attain at least perspectival self-consciousness. If he succeeds in such enquiries, he will end up knowing of other concepts that he falls under them, where these concepts are not anchored in the subject, and whose application has no special constitutive relation to what makes something his body.

Another sort of inquiry a thinker may have about his acceptance of such a content *I'm φ* is: is it reasonable to accept it? How is it that I have come to accept it? Since reason-giving states and events are characteristically conscious states and events, scrutiny and assessment of one's own reasons requires reflective self-consciousness.

I remarked earlier that a subject can be perspectival self-conscious without having the capacity to think of his own mental states at all. Reflective self-consciousness certainly does require that capacity, so perspectival self-consciousness does not imply reflective self-consciousness. I suspect that in principle reflective self-consciousness can exist without perspectival self-consciousness, though arguing that involves commitments on some heavy-duty theses about what the ability to represent one's own mental states involves. Without wading into those issues at this point, what seems to me less contentious is the existence of various kinds of enquiries which, in the nature of the case, require the cooperation of both reflective and perspectival self-consciousness. I mention two kinds.

The first kind concerns, broadly, a subject's understanding, explanation and correction of his particular conception of his situation in the world and his relation to it.

A full understanding of why one perceives the world as one does will involve a comparison of one's own case with how others perceive the world, when they are situated in a certain way in the world. One wide range of representations of how others are situated in the world will involve the subject's appreciation that the others fall under certain concepts not anchored in the subject. For the subject to gain any understanding of his own situation from this appreciation of the situation of others, he needs to appreciate that he too falls under those same concepts not anchored in the subject – and that is part of what is involved in perspectival self-consciousness. This appreciation then puts him in the position to consider possible errors to which he may be prone, having appreciated that others are vulnerable to them. And in the part of the assessment that involves consideration of his own mental states, either for explanation or for correction, the subject has to exercise his capacity for reflective self-consciousness. Each

kind of consciousness is indispensable in such a project. Without the perspectival self-consciousness, the subject would lack the ability to see the relevance of others' situation to his own. Without reflective self-consciousness, the subject would not have reflective knowledge of the states and events whose explanation and possible correction is in question.

Since understanding and explanation of one's own cognitive situation is part of what Bernard Williams calls the absolute conception, it is a corollary of this point that both perspectival self-consciousness and reflective self-consciousness are necessary components of any elaboration and application of the absolute conception.²⁷

A second kind of case in which perspectival and reflective self-consciousness cooperate involves a class of cases in which one comes to know that another person has a certain conscious state – be it a sensation, a perceptual experience, an emotion, a conscious thought – because one knows that one has it oneself in circumstances of the same type. This kind of case is obviously particularly relevant to the application of the techniques of attribution emphasized in simulationist approaches to the mental. In a range of cases within this general class of attributions, one can make the attribution to the other only if one appreciates that the other is in the same circumstances as oneself. This will involve attribution to oneself of properties that one attributes to the other, in the third-personal case (bodily properties, in the case of the attribution of sensations and some perceptual experiences). So here one needs reflective self-consciousness to be at the starting point of ascribing the conscious states to oneself; and one needs perspectival self-consciousness to grasp that oneself and the other are in the same kind of situation, and so makes the psychological ascription of the conscious state-type to the other reasonable.

The same point applies (independently of a general commitment to simulationism) to any concept of a mental state or event, where the nature of the concept gives a privileged place to first-person application. For such a concept, knowledgeable application to others involves both perspectival and reflective self-consciousness.

The interest of reflective self-consciousness is of course by no means restricted to such epistemological concerns. Since subjects are, in the first instance, bearers of conscious states (that is what makes them subjects), the initial basis of evidence for any psychological hypothesis about further, hidden properties of the subject will be the character of these conscious states. Access to that initial basis in one's own case again requires reflective self-consciousness.

27. B. WILLIAMS, *Descartes: The Project of Pure Enquiry* (Harmondsworth: Penguin, 1978): see the index.

VII. WHAT IS THE RELATION BETWEEN REFLECTIVE SELF-CONSCIOUSNESS AND A CONCEPTION OF MANY MINDS?

I suggest that the direction of philosophical or constitutive explanation of understanding runs from reflective self-consciousness to a mastery of the conception of many subjects of consciousness. Each of us has a general notion of a subject of consciousness that involves the idea that such a subject is something of the same general (Aristotelian) kind as he himself instantiates. Each of us has to work out what is involved in being of that kind, of what relations one has to have to things, events and states to be of that kind. Exercising reflective self-consciousness is a crucial part of working what is involved in being of that general kind, a subject of consciousness. So on this view, not only notions of conscious states and events are ones a thinker grasps from his own case, and in which the first person plays a special role. The same is true of the general concept of being a subject of consciousness, grasp of which also, though in a different way, involves a special relation to exercise of the first person.

On this approach, the general concept of being a subject of consciousness is another example of what I called relation-based understanding.²⁸ What makes thought thought about subjects is that they are conceived of as having a certain relation to the thing one thinks about in the first person way. As in other cases of relation-based thought, this means that we can make sense of the idea that someone else is still using the same concept, provided it is constrained by that relation, even though this other person has radically mistaken beliefs about what subjects are like. I think that Descartes and we are still thinking about subjects of consciousness, and thinking about them as subjects, even though he and we radically disagree about whether they have to be material, whether they are immortal, and so forth.

This treatment contradicts any thesis that implies that the first person can be used in thought only if the thinker already has a conception of many minds. So much needs to be done to argue against those theses. But I would like to close this section by emphasizing how what I am offering is, in its nature, an approach under which being a subject of consciousness makes available a uniform conception of multiple subjects. At the level of thought, the idea of multiple points of view is founded on the idea of an individual point of view. The other points of view are conceived of as things of the same kind as one's

28. 'Relation-Based Thought, Objectivity, and Disagreement', *dialectica* 64 (2010), 35-56.

own. This is an element of subjectivity in the idea of multiple subjects of consciousness. But there is also a uniformity across all thinkers' conception of multiple subjects of consciousness. The crucial point is that there is a level of description, corresponding to the level of description of indexical types in the theory of indexical thought, at which it is the *same* kind of subjectivity for each thinker. Each thinker x conceives of other points of view as things of the same type as he ([self] _{x} in the theory of indexical senses) enjoys.²⁹ In this conception that each has, the type of first-person thought – the [self] type – is uniform. It is also uniformly conceived as the type it is.³⁰ So there is a common element in the subjectivity. When we abstract from the particular thinker who is doing the thinking, the kind of conception is the same across different individual thinkers. The position is not one of a kind of solipsism of thought.

VIII. CONCLUDING REMARKS

This has not been a paper about the metaphysics of being a subject of consciousness. Here I have been more concerned with the epistemic and the conceptual, rather than with the metaphysics. But it is striking at how many points we need to appeal to the metaphysics of a subject-matter to explain adequately epistemic and conceptual phenomena. Three such points are salient:

(1) We needed to appeal to the metaphysics of the ownership of conscious events and states in explaining the status as knowledge of the self-ascriptions made in reflective self-consciousness.

(2) We needed to appeal to the metaphysics of what makes one body rather than another a particular subject's body in explaining why certain perceptual judgements were merely basic uses of the first person, rather than being exercises of perspectival self-consciousness.

(3) It is plausible that certain concepts are anchored in the subject only because the metaphysics of what it is to be the property that is picked out by such a concept itself has to do with the nature of subjects and conscious properties.³¹

I suspect these three points may be an illustration of a general lesson that applies far beyond issues of self-consciousness: the lesson that the metaphysics

29. '[self] _{x} ' is the notation of 'Demonstrative Thought and Psychological Explanation': see that paper for further discussion.

30. For the enthusiasts for the theory of senses: the [self] type is conceived under its canonical mode of presentation. See Chapter 8 of *Truly Understood*.

31. See 'Subjects and Consciousness'.

of a domain is sometimes a resource on which we need to draw to explain features of thought, knowledge and the forms of consciousness.³²

Christopher PEACOCKE
Columbia University
University College London

32. Earlier versions of this material were presented in 2009 in seminars and colloquia at Columbia University, University College London, and the University of Texas at Austin. More recent versions formed the basis of my Gareth Evans Memorial Lecture at Oxford University in March 2010, and a presentation at a conference in honour of Tyler Burge at UC Santa Barbara in May 2010. Amongst the many helpful comments on these occasions, I remember particularly those of Tim Bayne, Karen Bennett, Ned Block, Tyler Burge, Alex Byrne, Krista Lawlor, Rory Madden, Michael Martin, Lucy O'Brien, Jim Pryor, Michael Rescorla, Mark Sainsbury, Michael Tye and Ralph Wedgwood. I thank Fred Neuhauser for expert guidance through the continental literature, and Béatrice Longuenesse for her substantive comments on an earlier draft, for many valuable editorial improvements – and indeed for originally proposing that I connect my approach to these issues with the complex and interesting continental literature. Limitations of space have precluded the longer discussion of these connections that they merit.