

## Interpersonal self-consciousness

**Christopher Peacocke**

© Springer Science+Business Media Dordrecht 2013

If one were to write a book titled *The Varieties of Self-Consciousness*, one would start off with some distinctions. It will help to locate my topic in relation to those distinctions.

The first distinction concerns that kind of self-consciousness which involves only the minimal ability on the part of a subject to self-represent, to be in mental states with first person content, be it conceptual or nonconceptual. This minimal ability involves very little as compared with the more sophisticated states of which humans are capable. First person content can be present in the perceptual states of a creature capable of representing only some of its physical states and its relations to the physical environment, and not capable of representing anything beyond that. So the first distinction is between this minimal case, and richer varieties of self-consciousness. Some writers use 'self-conscious' to apply to a subject who is in any mental state involving first person content, and so they would apply it even in this minimal case. If that is done simply for terminological convenience, it is unproblematic. The terminology cannot, however, settle the question of whether there are distinctive kinds of mental states with first person content that in one way or another go beyond the minimal conditions for enjoying states with first person content. There are in fact several kinds going beyond the minimal.

One variety of self-consciousness plausibly going beyond the minimal conditions for first person content involves distinctive knowledge that one is in a certain kind of conscious mental state, or enjoying a certain sort of conscious mental event. This is sometimes called reflective self-consciousness. A second variety going beyond the minimal is the distinctive kind of self-consciousness for which passing the mirror

---

C. Peacocke (✉)  
Columbia University, New York, NY, USA  
e-mail: cp2161@columbia.edu

C. Peacocke  
University College, London, UK

test, if properly carried out, can provide evidence.<sup>1</sup> It has proved hard, and of some philosophical interest, to characterize that kind of self-consciousness. Intuitively it involves the ability to come to know things about oneself in ways that are not fundamentally tied to the first person, ways in which one can come to know things about other subjects too. This variety of self-consciousness I call ‘perspectival self-consciousness’, and it is distinct from reflective self-consciousness.<sup>2</sup> My topic here is a third variety of self-consciousness going beyond the minimal case, a variety which I call interpersonal self-consciousness. Interpersonal self-consciousness is awareness that one features, oneself, in another person’s consciousness.

Interpersonal self-consciousness, in one or another of its many subvarieties, is a ubiquitous feature of human life with its many social interactions. Awareness that one features, oneself, in another person’s consciousness is present in ordinary face-to-face conversation with another person, in e-mail exchanges, in the nonverbal interaction with another driver as one looks at him to resolve who is to cross the intersection first, and in nonlinguistic interactions between a mother and a young child. As with the other varieties of self-consciousness, interpersonal self-consciousness raises a number of constitutive, psychological, and epistemological issues.

My aims in this paper are to characterize more sharply a particular kind of interpersonal self-consciousness, which for reasons to be given later I will call ascriptive interpersonal self-consciousness; to describe the involvement of such interpersonal self-consciousness in a range of psychological, social and linguistic phenomena; and to address some of the theoretical issues surrounding interpersonal self-consciousness. These issues include the way in which knowledge that involves interpersonal self-consciousness can be attained; the role of interpersonal self-consciousness in perception and in various theories of understanding; its relation to the other varieties of self-consciousness; and the status of some of the philosophical theses about interpersonal self-consciousness found in the literature. Though I will not be discussing the history of philosophy here, interpersonal self-consciousness is also a foundational notion for the now centuries-long discussion of recognition of one subject by another in German philosophy. It is equally a pivotal notion for a tradition in sociology dealing with the self as represented by others, in writers from Mead (1934) to Erving Goffman (1959).

Interpersonal self-consciousness is much closer to what is meant by ‘self-consciousness’ in everyday conversation than are some of the other notions discussed by philosophers past and present under the heading of self-consciousness, reflective and perspectival self-consciousness included. Some of the significance of interpersonal self-consciousness is the significance of that everyday notion. But

<sup>1</sup> Two landmarks in the empirical literature: Gallup (1970) and Reiss and Marino (2001).

<sup>2</sup> There is further discussion of reflective self-consciousness and perspectival self-consciousness, and of some constitutive and evidential issues surrounding the mirror test, in Peacocke (2010). That paper also contains a more explicit characterization of perspectival self-consciousness. I should add that on my view, first person content is found both in conceptual content and in non-conceptual content. For further discussion, see my paper ‘Subjects and Consciousness’ (Peacocke 2012).

before we can argue for that or anything else about the notion, our first task must be to say more precisely what interpersonal self-consciousness is.

## 1 Illustration and delineation

A soldier is on patrol in Afghanistan. He is a member of a group, walking at night, in single file along a mountain path. The group proceeds as quietly as possible, to avoid detection by the enemy. Our soldier suddenly hears a click, just a few feet away. He freezes. It is the click made by a rifle being switched from the safe position into one ready to fire.

Our soldier freezes because he is instantly aware that he is the object of someone else's consciousness. Given all the circumstances, he is aware that the person in whose consciousness he features will think of him as a soldier, hence a person, someone capable of what we, though not necessarily he, will call first person intentional states. Both we, and the person holding the rifle that made the audible click, will distinguish between the case in which he merely perceives an object on the path, and perceives a person. So will the person perceived. What the soldier on the path wants to avoid is the case in which he is represented in the other's consciousness as a person.

More generally, there are at least three broadly different ways in which it can be true that one is aware that one features, oneself, in another's consciousness. The three ways are distinguished by the degree of richness of the conception of oneself that one is attributing to the other person. One may be aware that the other person is representing oneself merely as an entity in the other's environment, without the other person even recognizing that one is a conscious subject. The rifle-holder in our example may not have distinguished you from some movement in the foliage next to the mountain path. In a second kind of case, you may be aware that the other is representing you as a subject of mental states, but not one capable of first person propositional attitudes. In a case of the third kind, you are aware that you feature in the other person's consciousness as a self-representing subject, as a person capable of first person mental states about yourself. These are cases of consciousness of yourself being represented by another person as a self-representing subject. It is concern about being an instance of this case that makes our soldier freeze. The case is not confined to the dramas of conflict. The case is instantiated myriad times in our ordinary conversational and non-linguistic interactions with one another. This is the kind of interpersonal self-consciousness, awareness that one features as a self-representing creature in another's consciousness, that I aim to investigate here.

It is a defining feature of this species of interpersonal self-consciousness that the first person notion or concept plays a double role in its instances. The subject who is interpersonally self-conscious with respect to another is aware of himself as represented in a certain way in the other's consciousness. This awareness about himself as so represented has a first person content. It is awareness with a content of the form "That other person is conscious of *me* in such-and-such a way". But the relevant "such-and-such way" also involves the first person too. The content "conscious of *me* in such-and-such a way" is not merely consciousness of me as a

mere material object. The particular way in which our subject thinks of the other as representing him is one that involves the other ascribing the capacity to enjoy states with first person contents to our subject. This ascription on the part of the other involves reference to the first person notion or concept. The interpersonally self-conscious subject enjoys an awareness that the other is referring mentally to the first person notion or concept, and is ascribing the capacity to use it to the subject himself. Because of this distinctive second role of the first person in such cases, I call this species of interpersonal self-consciousness *ascriptive* interpersonal self-consciousness. Ascriptive interpersonal self-consciousness is awareness that you are being represented by another person as a self-representing subject.

The description of the mental states involved in ascriptive interpersonal self-consciousness trips off the tongue in ordinary English; but in fact the way in which these states involve the first person concept involves considerable complexity in the theory of intentional content. Suppose that you are ascriptively interpersonally self-conscious with respect to me. Then you are aware that I am not merely employing, I am also referring to the first person concept in my attribution of mental states to you. I am attributing to you use of the first person concept. So—already simplifying somewhat—I am referring to the first person concept (actually under its canonical concept), and thinking that you are employing it.<sup>3</sup> So I am employing a concept of a concept. Since you are aware that I am doing this—you are aware that I am employing that concept of a concept—we are now at a third level in the Fregean hierarchy of concepts (senses). You are thinking about the concept of a first person concept that I am employing. In doing so, you are using a concept (actually once again, the canonical concept) of a concept of a concept. For those interested in such matters, I include a formal treatment of the case, using neo-Fregean apparatus, in the [Appendix](#).

The example of the soldier may be vivid, but it differs from more ordinary cases of ascriptive interpersonal self-consciousness in several respects. In everyday cases, such interpersonal self-consciousness is symmetrical: each person is interpersonally self-conscious with respect to the other. In the everyday cases, each person perceives the other. Each person's interpersonal self-consciousness involves the other person as given in perception, rather than in some mixed descriptive-demonstrative such as “the person whose rifle made that click”.

Certain examples of ascriptive interpersonal self-consciousness can arise simply from mutual perception when it attains a certain degree of embedding. It can be helpful for various purposes to develop a description of these common examples, and I will do that briefly now. The description to be developed is a description of an attained state of interpersonal self-consciousness, rather than a description of how the state might be attained. We will come to the ‘how’ question in a later section.

So, let us take a case in which two subjects see each other. The example is not the only instance of this perceptual subspecies—it could not be so, because the blind can also certainly enjoy states of interpersonal self-consciousness. But the visual

<sup>3</sup> The simplification consists in the fact that I am referring to the first person type; what you are employing in your thought is an instance of that type. More on this below.

case is easier (at least for the sighted) to think about in relation to degrees of embedding. We have then two subjects  $x$  and  $y$ , and each sees the other:

- (1)  $x$  sees  $y$
- (2)  $y$  sees  $x$ .

It may aid in thinking about the case to imagine yourself as  $x$ , and some other subject as  $y$ , and I will sometimes expound features of the case using such an identification. Conditions (1) and (2) are jointly so weak that they are consistent with each thinking the other does not see him. At the level of conditions (1) and (2), we have no embedding of seeings in other seeings.

To reach the level of the First Embedding of mental states, we add to the description of the case that

- (3)  $x$  sees that  $y$  sees him
- (4)  $y$  sees that  $x$  sees him.

If you are  $x$ , then when (3) is true, you might express this fact by saying of  $y$ , “He sees me”.

The First Embedding conditions (3) and (4) are, however, entirely consistent with each of  $x$  (you) and  $y$  (the other) featuring merely as an object in the other’s awareness. You and the other may each think that he is being seen only as a material object, and not as a conscious subject.

At the level of the Second Embedding of mental states, we then add to the description of the case that

- (5)  $x$  is aware that (4)
- (6)  $y$  is aware that (3).

When (5) and (6) hold, you and the other must each feature as a subject in the other’s consciousness, under the presumption that seeings can be attributed only to subjects. When (5) holds, and you are  $x$ , you are aware that the other sees that you see him. What you are aware of is a state of affairs in which the other sees something to be the case that involves your being a subject (an involvement of which you can presume the other has some kind of appreciation).

The fact that these Second Embedding conditions hold does not yet guarantee ascriptive interpersonal self-consciousness in the sense identified above. The Second Embedding conditions ensure that you are aware that the other sees that you see him. What you are aware of is that the other enjoys a seeing with the content something like (7), as thought by the other person:

- (7) That person sees me

where, since this is thought by the other, that last “me” refers to the other. Here you are indeed ascribing states with first person contents to the other; but you are not thereby representing him as doing the same vis-à-vis you. That is, you are not yet attributing to the other person a mental state which involves ascription to *you* of the use of the first person notion or concept.

Similarly, at this level of Second Embedding, what the other is aware of is that you enjoy a seeing with a content something like (8), as thought by you:

(8) He sees me.

The other is here ascribing states with first person content to you; but he is not yet attributing to you a mental state which involves ascription to *him* of use of the first person notion or concept.

We do not reach the perceptual subspecies of ascriptive interpersonal self-consciousness until the Third Embedding conditions are fulfilled. The Third Embedding conditions are probably most easily grasped when framed in terms of “you” and “the other”. When the Third Embedding conditions hold, each person is aware of the other’s awareness as described in the Second Embedding conditions. A simple, and near-enough accurate, way to formulate an instance of this Third Embedding, without involving barely graspable iterations of “sees”, is (9):

(9) You are aware the other is aware you’d sincerely say “He sees me”.

Now in (9), what you are aware of is something that involves the other’s appreciation that you are employing the first person notion or concept. The other’s attribution to you of grasp of the first person features in your own awareness. So here we do have a genuine case of ascriptive interpersonal self-consciousness.<sup>4</sup> Good manners require relegation of the formalities of the situation to a footnote.<sup>5</sup>

Distinctions drawn at these various levels of embedding can contribute differentially to the explanation of action. Take first the distinction between no embedding and First Embedding. When a spy sees that you see him, he will put away or hide his telescope to avoid detection of his activity. No such action is explained merely by the fact that you see him, or that he sees you. What about the additional explanatory power of the Second Embedding relative to the First? You see that he sees you; but this is consistent with his not knowing that you do. So actions of his that can, in context, be explained only by such knowledge will be explained with the Second

<sup>4</sup> A few sentences back, I wrote “near-enough accurate”, because it should also be added that the sincere saying “He sees me” must be suitably based on the subject’s own visual experience, rather than knowledge from inference or testimony.

<sup>5</sup> When (9) holds, what you are aware of is a state of affairs in which the other has an awareness with a content:

(i) That person sees that I see him.

How should this be regimented in neo-Fregean terms? Let  $\langle \dots \rangle$  be the sense of ...; let [self] be the first person type of sense. [self]<sub>x</sub> is the instance of the first person type usable only by *x* to think of himself.  $\wedge$  is predicational combination in the realm of senses. (The looseness in these characterizations can easily be amended by anyone who takes notice of it.) This is the terminology and apparatus of Peacocke (1981). We can then regiment the content expressed by (i) in neo-Fregean terms thus, as thought by the other: for some mode of presentation *m*, the mode of presentation *m* under which that person sees me,

(ii) Sees-that (that person,  $\langle \text{see} \rangle^m \wedge [\text{self}]_{\text{that person}} \rangle$ ).

When you are aware that (ii) is the content of the other’s awareness, the condition for ascriptive interpersonal self-consciousness is met. That is so because the content (ii), which specifies the content of the other’s awareness, of which you are aware, is one that attributes to you use of the first person. This is a more formal way of explicating how Third Level embedding conditions imply ascriptive interpersonal self-consciousness.

Embedding conditions being fulfilled, but will not be explained by First Embedding conditions. An example: when he knows that you see that he sees you, he can take something you've visibly done as an indication that you intend to communicate something, (say) that you will attend a certain event. If he didn't know that you see that he sees him, he wouldn't have reason to do that; and so forth.

Someone may protest that the alleged distinction between the Second and Third Embedding is of no significance for the present topic, on the ground that anything that is a conscious and seeing subject must have the capacity for some kind of first person representation, must be capable of mental states with first person contents, whether conceptual or nonconceptual. I dispute the cited ground. It seems to me that we can conceive of conscious subjects who represent objects, events, and their properties and relations, what is in fact their environment, without use of the first person. Such subjects employ a *here* and a *now*, but not an *I*. These are the subjects at what I called 'degree 0' of self-representation in their conception of the objective world in 'Subjects and Consciousness' (Peacocke 2012). If you disagree about the possibility of such cases, that does not matter for the purposes of the present paper. It means only that the conditions for ascriptive interpersonal self-consciousness can be met sooner in the hierarchy of embeddings than on the opposed view.

The ascriptive interpersonal self-consciousness on which I have been focusing is manifestly only one of many notions of interpersonal self-consciousness that we can identify once we have a hierarchy of attitudes and of concepts of concepts. Its importance is psychological, in at least two connected respects. First, in appreciating that another is aware of oneself as a possessor of first person attitudes, one is appreciating that the other is aware of oneself as one really is, with the capacity for self-representation one really has. This is a kind of recognition. Second, this kind of recognition is enormously important to us in our personal, social, and political relations, and in our emotional lives too. This recognition is a precondition of close personal relations, and it has analogues in social and political relations. The point is already present in William James, who took interpersonal self-consciousness as a subvariety of self-consciousness in his chapter 'Consciousness of Self' in *The Principles of Psychology* (1918). He wrote, "A man's Social Self is the recognition which he gets from his mates" (p. 293); and "No more fiendish punishment could be devised, were such a thing physically possible, than that one should be turned loose in society and remain absolutely unnoticed by all the members thereof" (p. 293). Solitary confinement is regarded with fear by even the hardest of prisoners (Abbott 1981).

## 2 Some roles of interpersonal self-consciousness

Interpersonal self-consciousness is inextricably implicated in a range of social and psychological phenomena. Here are five examples.

### 2.1 Joint action

In any situations of joint action beyond the more primitive cases, and certainly in any ordinary situation of joint action of two humans, each participant will be

interpersonally self-conscious with respect to the others. Suppose you and I are engaged in some joint action. This will involve my awareness that you know about my first person knowledge and beliefs. This awareness of your knowledge of me affects the way I act in carrying out our joint project. The same applies to you vis-à-vis me. You will on occasion draw on your awareness of my knowledge about your first person knowledge and belief about yourself. We would make mistakes in joint action if there were no such awareness. Maybe we are dancing together. Knowing that you will not dance well if we try to make moves we are not both confident we can make, I will not lead you into a move that I know you believe you can't reliably make.

The interpersonal self-consciousness required in successful joint action has a social dimension that is, *prima facie*, quite different from two of the other varieties of self-consciousness I mentioned at the outset. One of those was reflective self-consciousness, awareness of one's own mental states (as one's own), awareness attained in a distinctive way. Reflective self-consciousness by itself is of no help in joint action unless the other also has some awareness of those states, which is what ascriptive interpersonal self-consciousness involves. The same applies to perspectival self-consciousness. Perspectival self-consciousness does involve some kind of conception of oneself as an object in the world of the same kind as other objects. But just as in the case of awareness of one's own mental states, this perspectival self-consciousness will be of no help in joint action unless it is also connected with another subject's knowledge of oneself.

It may be that there is a philosophical case to be made that there is some connection between one or another apparently non-social notion of self-consciousness and interpersonal self-consciousness. The point at the moment is just that the case would need to be made. Perspectival self-consciousness and reflective consciousness of one's own mental states are not apparently themselves social notions.

## 2.2 Linguistic communication

Linguistic communication is a special case of joint action and joint awareness. This is so whether communication is face-to-face or in any other medium. In any such communication, at the starting point, there may be more or less knowledge of your interlocutor's history and circumstances. But as communicative exchanges proceed, an occasion-specific background is built up of what has been said and possibly accepted, what denied or disputed, what knowledge has been revealed (and what has not). This accumulating occasion-specific background helps to determine the nature of the joint attention in such a conversation. Attitudes that involve interpersonal self-consciousness will be essential in specifying the nature of this accumulating background. You may for instance be aware that your interlocutor knows that you represented yourself as knowing so-and-so, for instance. This affects the joint perception of the conversation, and how it proceeds. Human conversation would be utterly different if it did not involve multiple instances of such interpersonal self-consciousness: it would hardly be recognizable as normal conversation at all.

It is also the case that specification of success in a linguistic exchange, an aim that is essential to the nature of linguistic communication, already involves

interpersonal self-consciousness. When you say that  $p$ , and I understand you, in central cases I am aware that you know that I am aware you have asserted that  $p$  to me. This awareness of mine is an instance of ascriptive interpersonal self-consciousness.

### 2.3 Understanding the second person

The second person pronoun is a word whose successful use involves ascriptive interpersonal self-consciousness. Suppose you say to me, “You are F”. Your intention in the primary, central case is that I should thereby come to know that I am F. I understand your utterance only if I come to know that you are saying that I am F. As many writers have noted, it does not suffice for understanding a second person utterance that I know that you intend, of whoever is your addressee, that that person come to know that he is F. I understand your utterance only if I come to know that I am your addressee, and that what you are saying is that I am F. In a case of your successful communication with me in which you use the second person, I as audience know that you, the speaker, are aware that I know that you are saying that I am F. This is, in more than one way, an instance of ascriptive interpersonal self-consciousness on my part. One source of the interpersonal self-consciousness here is the very first person content that I know you are saying to me. A successful practice of using the second person is possible only for creatures who are interpersonally self-conscious with respect to one another.

This description of what is involved in using and understanding the second person does not invoke a special second person concept or way of thinking. It uses only third person and first person singular concepts, and concepts of those concepts (and further concepts thereof, up the Fregean hierarchy). So as not to distract attention from the description of the range of cases in which interpersonal self-consciousness is involved, I defer consideration of whether there is a second person concept, not explicable in terms of the first or third person concepts, to Sect. 3.

### 2.4 ‘Self-conscious’ in ordinary usage

Here are four quotations selected by the editors of the *Oxford English Dictionary* (Second Edition): from Carlyle, in *The French Revolution*: “Self-conscious, conscious of a world looking on”; from J. Hawthorne, *Fortune’s Fool*: “He was singularly free from self-consciousness; and she was so exquisitely self-conscious as completely to conceal it”; from Orwell’s *Road to Wigan Pier*: “Self-conscious Socialists dutifully addressing one another as ‘Comrade’”; and from J. Irving, *158-Pound Marriage*, “She had never been self-conscious about what she wore”. Lucy O’Brien (2012) has also discussed very well the ordinary notion of self-consciousness expressed in these sentences. The ordinary notion of being or feeling self-conscious involves ascriptive interpersonal self-consciousness in the sense I have been trying to delineate. Feeling self-conscious involves awareness that there is a certain standard others think your actions should meet, a standard they think you should know, and which you are concerned to meet (or they think you should be). That is an awareness that others are attributing to you certain first person attitudes: knowledge that such-

and-such standards are the ones you should be meeting. Ascriptive interpersonal self-consciousness is a precisification and generalization of a notion involved in our everyday applications of a notion of being, or feeling, self-conscious.

## 2.5 Relationships of mutual concern

In a relationship of mutual concern, one person may do something for another because the first realizes that the second has a certain attitude, including an emotion, with a first person content. The second person may realize that this is the reason the first person is doing it. This realization involves an exercise of interpersonal self-consciousness. The first person may also value this realization on the part of the second person. Relations based on anything from mere recognition of common humanity with a neighbor, to friendship and up to and including deeper relationships, may thus essentially involve exercises of interpersonal self-consciousness. Such capacities for interpersonal self-consciousness, and for their recognition, can also of course be abused by the cruel person.

## 3 Is there an irreducible second person concept?

It is, or ought to be, uncontroversial that there is a use of “you” in thought, and that it can occur in thinking that is not expressed out loud in language, and is not intended as part of any communicative act. I may see someone driving dangerously, and think “If you go on driving like that, you will be involved in an accident”. One of the philosophical questions in this area is not whether there is such a use of “you” in thought, but whether it, or anything else, requires us to acknowledge the existence of a second person concept that is distinct both from third person concepts and first person concepts, and concepts explained in terms of such concepts. It certainly seems that in the example just given, the use of “you” in thinking involves imagining addressing the driver in question as “you”, an expression we have explained in (ii) as not involving an irreducible second person concept. The driver thus imagined as the addressee is thought of as *that driver*, a perceptually based mode of presentation. The thesis that there is no distinctive second person concept is equivalent to the proposition that any occurrence of “you” in thought can equally be explained in terms of other concepts that do not involve an irreducible second person concept. I myself am in agreement with Richard Heck, when he writes, “there is no such thing as a second-person belief” (2002, p. 12).<sup>6</sup>

A thinker may use “you” in thought without realizing that he is referring to himself. Take John Perry’s example in which he is pushing a trolley in a

---

<sup>6</sup> Heck also says that the claim is obvious, and writes, “I don’t really know how to argue for this claim”. Presumably what has to be argued for is the eliminability claim in the text above, a claim that can be made increasingly plausible as explanation of uses of “you” in terms of other concepts come to seem generalizable across examples. Sebastian Rödl, in his book *Self-Consciousness* (2007, p. 187), legitimately complains against Heck’s remark that “you” can be analyzed as “That person to whom I am speaking”; but that point does not establish an ineliminability thesis.

supermarket, seeing someone in a mirror at an angle above pushing a trolley from which a stream of sugar is pouring on to the floor (Perry 1979). Perry may think “You are making a terrible mess”. He might even shout out loud “You in the grocery aisle, you’re making a terrible mess”, without realizing he himself is in the grocery aisle. But Perry does not think “I am making a terrible mess”. If the notions of concept and mode of presentation are to be constrained by their role in cognition in the most basic way, this shows that the concept “you” employed by Perry on this occasion and the first person as employed by Perry are distinct. This conclusion seems to me to make it impossible to accept the thesis of the penultimate sentence of Sebastian Rödl’s book, which states that “Second person thought is first person thought” (2007, p. 197).

Rödl raises a series of objections which would apply against what I have said about understanding the second person pronoun. In discussing McDowell, Rödl writes,

...the capacity to know that thoughts stand in a certain relation depends on the capacity to share them. Understanding you, who are addressing me with “You...”, McDowell proposes, I know that your thought stands in a certain relation to a thought I think and which I would express by “I...”. But how do I represent your thought in knowing this? If I do not think your thought, how then does it figure in my thinking? Figure there it must, if I am to know that are you thinking it (p. 196).

Rödl concludes the paragraph from which the preceding quotations are taken with the claim that “I possess the notion of a thought that bears the relevant relation to my thought [a thought I express by “I...”—CP] only if I understand you; my understanding you is the source of my possessing that notion, not the other way around” (p. 196).

So let us consider a case in which *a* is the utterer of “You are F”, and *b* is the hearer who understands this utterance, which is in fact addressed to him. In a central case, *a* intends *b* to gain the information that *b* would express in a sincere utterance of “I am F”. On one neo-Fregean treatment of indexicals, this is the Thought  $[\text{self}]_b^{\langle \text{is F} \rangle}$ , where  $\langle \dots \rangle$  denotes the sense of ... and  $[\text{self}]_b$  is the first person type indexed with the object *b*. For more on this apparatus and its deployment, see Peacocke (1981). In a second person utterance, the utterer does mean the hearer to convey first person information to the hearer, and means the hearer to appreciate that. Rödl’s question can then be raised about how *b* represents *a*’s intention, and knowledge, about what information *a* is providing to *b* in this second person utterance. But Rödl’s question does have a straightforward answer on a neo-Fregean treatment of indexical content. The utterer *a* refers to the first person type of mode of presentation, and intends, of the instance of it that *b* is capable of employing in thought (viz.  $[\text{self}]_b$ ), that his hearer come to know the Thought that consists of it combined with the concept  $\langle \text{is F} \rangle$  in predicational combination. So I, like any neo-Fregean who proceeds in this way, reject Rödl’s principle that “the capacity to know that thoughts stand in a certain relation depends on the capacity to share them”. You can know that my first person Thought is of the same first person type as your first person thoughts without having attitudes to the very same Thought.

In this respect, first person Thoughts and present tense Thoughts should be treated on a par. You and I can know now that Napoleon at noon on 1 January 1810 thought (let us suppose he wrote it in his private diary) “Now would be a good time to invade England”. We know that Napoleon’s Thought was of the present tense type. Only those thinking at noon on 1 January 1810 can think of that time in the present tense way, can think of it using a mode of presentation of the [now] type, as neo-Fregeans would say.

Rödl himself thinks the temporal parallel cuts the other way, saying that ““Today...” said yesterday and “Yesterday...” said today express the same act of thinking” (p. 196). The token acts of a thinking occurring today and a thinking occurring yesterday are distinct, since they occur at different times. Rödl’s claim is presumably that they have the same Fregean content (so we could say that that the utterances in question express the same act-type). I am not disputing that there is some special relation between the Thoughts expressed by “now” as uttered yesterday and “yesterday” as uttered today. The latter thought is how the earlier one must be updated in one’s memory if information is to be stored in a form that preserves correctness (and a fortiori knowledge). But that does not make them identical Thoughts. I will not proceed with a full-dress critique of the claim that they are identical, but will make two observations.

First, if the claimed identity holds, what are we to make of the Thoughts “Five minutes ago it was raining” thought now, “300 seconds ago it was raining” thought now, and “It is raining” thought five minutes (equals 300 seconds) ago? If the claimed identity holds, then the first and second Thoughts specified here are identical with the third. Yet someone could rationally judge the first without thereby judging the second, if she did not know how many seconds there are in a minute (or how many seconds there are in five minutes, for that matter).

The other observation is that identity of sense in Frege is answerable to facts about cognitive significance at a given time. There certainly will be constraints on what it is rational to accept as time passes. But those constraints can be formulated in terms of Thoughts that stand in a certain relation to one another, a relation that need not be identity.<sup>7</sup>

#### 4 Philosophical and psychological issues

How do we attain the awareness of others’ attitudes about us that is involved in interpersonal self-consciousness? An answer to this question must provide a means that is, on occasion, capable of yielding knowledge. Interpersonal self-consciousness is naturally characterized in terms of awareness-that, a state that is both factive and knowledge-entailing. We do often have such awareness.

One subclass of instances of interpersonal self-consciousness is provided by cases of full joint attention. Suppose I am seated next to a table; you and I are in the

<sup>7</sup> These are not the only points I would want to make about Rödl’s final chapter ‘The Second Person’. Since I have been critical on his treatment of “you”, let me also emphasize that there is much interesting and important material in the earlier chapters of his book.

same room; each of us is in the other's visual field, and so is the table. In an entirely normal case meeting these conditions, there is joint awareness between us that we are in such a situation. Simply for speed and smoothness of exposition, in describing this state of affairs, we can use Barwise's elegant formulation in terms of situations  $S$ , where (in a slight variant) I will write " $S: A$ " for "A holds in situation  $S$ " (Barwise 1988, 1989). Then on Barwise's treatment, the case just described is one in which there is a situation  $S$  meeting these conditions:

- S: I am seated next to a table
- S: I am aware of  $S$
- S: You are aware of  $S$ .

(This is a self-involving situation in Barwise's treatment. It may be possible to obtain the benefits of self-involvement in a different way, by speaking of states of joint awareness that make reference to themselves. Those interesting issues are not pivotal to the present topic.<sup>8</sup>) It has been emphasized in the literature that situations such as  $S$  can generate the distinctive features of joint awareness and of common or mutual knowledge.<sup>9</sup> Computational mechanisms that operate on specifications of the situation of which I and you are aware in this example will generate the familiar embeddings. When there is such a situation  $S$ , the following can be computed from the above specification of the situation, as Barwise emphasized:

- I am aware I am seated next to a table
- You are aware I am seated next to a table
- I know that you are aware that I am seated next to a table
- You know that I am aware that I am seated next to a table
- I know that you know that I am aware that I am seated next to a table
- .
- .
- .

The fifth line here is an attribution of interpersonal self-consciousness, if the knowledge it attributes is conscious knowledge. It attributes to me knowledge that you are attributing to me an awareness with a first person content, that I am seated next to a table.

So that is one route to interpersonal self-consciousness. It is a kind of route that is instantiated many times over, every day, in our interactions with others, in respect of those parts of the worlds that we jointly perceive and with respect to which we are engaged in joint attention. The content "I am seated next to a table" could have been replaced with contents concerning my perceptible relations to anything else in the jointly perceived environment, or concerning my jointly perceived properties. Actually the example of iterated seeing and awareness that we considered earlier as an illustration of interpersonal self-consciousness was a special case of joint

<sup>8</sup> For further discussion, see the Appendix to Peacocke (2005).

<sup>9</sup> For the classic characterizations of common or mutual knowledge, see Lewis (1969) and Schiffer (1972).

attention, the case that is sometimes called ‘contact attention’. That was the special case of joint attention in which, again using Barwise’s apparatus, we have a situation *S* such that:

- S: I see you
- S: You see me
- S: I am aware of *S*
- S: You are aware of *S*.

In general, joint awareness may or may not have, as its most embedded content, a content that is first personal, or a content that is about seeing or perception. In visually based contact attention, the most embedded content has both of these features.

Joint awareness cannot however be the only source of interpersonal self-consciousness. There are cases of interpersonal self-consciousness in which you are aware of my knowledge and beliefs about myself that go far beyond anything that can be discerned from the currently jointly attended world, or from memories thereof. Because of your knowledge of my values, my history, my prejudices, my past encounters, our past interaction, I may be aware that you represent me as having attitudes about myself that relate to anything on this list. These attitudes of mine need not be available in joint awareness alone. Perhaps joint awareness is a basic kind of case of interpersonal self-consciousness that makes possible all the others. All the same, we still have to give some account of how we attain those instances of interpersonal self-consciousness that are not given in joint attention.

A very different source of interpersonal self-consciousness is second order simulation. A second order simulation is one in which you simulate someone else’s simulation. Second order simulation can yield the awareness involved in interpersonal self-consciousness. To attain interpersonal self-consciousness, it suffices for you to attain a state of awareness of, for example, this form:

- (10) You are aware: he believes you think you’re F.

How can you reach an awareness with the content following the semicolon in (10)?

You can reach it in three steps:

Step 1: You ascribe beliefs to him on the basis of simulation, and ask: what would you believe about yourself if you were in his situation vis-à-vis you?

Step 2: He knows my situation, my history, perceives my actions and their results; in that situation, if he also had the background attitudes ascribable to me on the basis of my actions, he would think that he is F.

Step 3: So by simulation, he believes that I think I’m F.

Second order simulation is here a mechanism for ascension through higher levels of embedding of mental states with content.

Interpersonal self-consciousness attained by this second order simulation need not be reached by conscious personal level inference. A subpersonal, content-involving computation can proceed through Steps 1 through 3. Correspondingly, the awareness attained need not be a judgement made by a sequence of conscious

rational steps. It may not be a personal level mental action at all. It can simply be an awareness that simply occurs to the thinker.

The need to work out, consciously and often with difficulty, what attributions to make to others, and what actions to perform, is something distinguishing those occupying some point on the autistic spectrum from those who are not autistic. The more autistic subjects have neither states of interpersonal joint awareness, nor subpersonal rapid simulation of other's mental states and likely actions. They have to reach some of the attributions we ordinarily obtain in cases by interpersonal self-consciousness by slow, conscious reasoning they often find painfully difficult. Here is a passage from Temple Grandin's description of her own autistic condition (2006):

When I have to deal with family relationships, when people are responding to each other with emotion rather than intellect, I need to have long discussions with friends who can serve as translators. I need help in understanding social behavior that is driven by complex feelings rather than logic.

Hans Asperger stated that normal children acquire social skills without being consciously aware because they learn by instinct. In people with autism, "Social adaptation has to proceed via intellect." Jim, the 27 year old autistic graduate student I have mentioned in previous chapters, made a similar observation. He stated that people with autism lack the basic instincts that make communication a natural process. Autistic children have to learn social skills systematically, the same way they learn their school lessons. Jim Sinclair summed it up when he said, "Social interactions involve things that most people know without having to learn about them." He himself had to ask many detailed questions about experiences other people were having to figure out how to respond appropriately. He describes how he had to work out a "separate translation code" for every new person (pp. 155–156).

Prima face, neither joint awareness nor second order simulation can be assimilated as a special case of the other. On the one hand, joint awareness does not seem to involve simulation of any kind at all. The iterations of knowledge or potential awareness do not involve considering what one would think or experience if one were in some other situation. The iterations flow from the actual structures and awareness, independently of any such consideration by the participants, and equally independently of running any belief-forming mechanisms off-line. On the other hand, in cases of second order simulation, the mechanisms that lift us to iterated attitudes seem to have nothing to do with any of the various distinctive properties that have been proposed as crucial to full overt joint awareness. There does not seem to be any essential involvement of self-involving situations, nor of states or events of awareness that make indexical reference to themselves, in successful second order simulation. Joint awareness and second order simulation seem to be psychologically independent ways of attaining interpersonal self-consciousness.

For second order simulation to be successful, it is important that it be available to the other person that one of your actions is, say, a reaching to a cup in order to drink. Without such information, he could not so successfully simulate you. He could not

so successfully use information about your past and current actions to attribute background attitudes to you. It follows that the existence of perspective-independent notions of particular actions, notions that can enter both the content of intentions and the content of another's perception of those actions as discussed in the literature on mirror neurons, plays an important role in facilitating interpersonal self-consciousness. Perhaps we can in principle conceive of the instantiation of interpersonal self-consciousness without such perspective-independent notions, but we could not attain it so easily or smoothly. Without the perspective-independent notions, there would have to be an additional layer of computational complexity, at which subpersonal mechanisms in a subject somehow compute basic intentional characterizations of actions from the physically described movements of another agent, drawing on who knows what further information. By contributing to the identification of another's goals, the perspective-independent notions contribute to the attribution of attitudes to another, and to facilitating interpersonal self-consciousness.

Vittorio Gallese has written of “a new conceptual tool: the shared manifold of intersubjectivity” (2005, p. 111). Gallese's shared manifold consists of specifications of events using perspective-independent notions that feature both in intention and in perception, specifications that I would put in the form “ $x$  is a  $\varphi$ -ing”, where  $\varphi$  is one of these perspective-independent notions. The  $\varphi$ -ing might be a drinking from a cup, or reaching for a particular object. Gallese writes, “in humans, and even more so in monkeys, the shared space coexists with but does not determine self-awareness and self-identity. The shared intentional space underpinned by the mirror matching mechanism is not meant to distinguish the agent from the observer”; “Rather, the shared space instantiated by the mirror neurons blends the interacting individuals within a shared implicit semantic content.” (p. 111) Gallese's thesis is that we recognize that other human beings are similar to us by means of this shared manifold, and that it makes possible “intersubjective communication, social imitation, and mind reading” (p. 115). In effect I have just been supporting Gallese's position, for the special case of interpersonal self-consciousness.

The approach I have been offering also suggests, however, that Gallese's shared manifold is just the first of a hierarchy of shared manifolds. The hierarchy is generated from Gallese's base manifold by the operation of applying a simulation mechanism to a given level, to attain attributions of mental states not represented at that lower level. The resulting attributions form a new level, to which simulation operations can be applied again, to generate another new level; and so forth. I suspect that the idea of a shared manifold of intersubjectivity has application at these higher levels too. As far as I can see, this extension is entirely within the spirit of Gallese's approach, and the extension has his shared manifold at its foundation, its basic level on which all other levels are built.

In a well-known paper ‘Five Kinds of Self-Knowledge’, Ulrich Neisser (1988) advances various claims that bear on the foundations of interpersonal self-consciousness. He introduces the notion of ‘the interpersonal self’, which is “the self as engaged in immediate unreflective social interaction with another person” (p. 41). Neisser's view is that this interpersonal self “can be directly perceived on the basis of objectively existing information” (p. 41). What is perceived in another is

“activity” (p. 41). In Neisser’s presentation of his points, this is offered as an extension of a Gibsonian conception to the intersubjective realm. I have already noted that the mental states involved in interpersonal self-consciousness need not be reached by conscious personal level inference. I think the convincing arguments of Fodor and Pylyshyn (1981) against the Gibsonian conception of information pickup as explaining perception apply straightforwardly in interpersonal perception. The arguments against the Gibsonian conception are entirely general, and not restricted only to some specific domain. The only argument offered by Neisser for ‘direct perception’ in his sense are studies about infant/mother interaction, and the distress suffered by infants when mothers’ interaction with their infants is not synchronized in the normal way.<sup>10</sup> This is not evidence for the Gibsonian conception, since the phenomenon is consistent with the infants’ perceptions of their mothers’ actions being a consequence of subpersonal computations. The computational structures underlying a computational process may of course be innate. But let us set aside that old, and (in my view) settled issue. For Neisser also propounds some more interesting claims that are independent of his Gibsonian allegiance, and that raise issues about the bases of interpersonal self-consciousness.

First, Neisser offers a characterization of intersubjectivity: “If the nature/direction/timing/intensity of one person’s actions mesh appropriately with the nature/direction/timing/intensity of the other’s, they have jointly created an instance of what is often called *intersubjectivity*.” (p. 41). There can, however, be meshing of such actions without intersubjectivity. Suppose there is what I take to be an inanimate robot, in motion nearby to me. I may delight in copying its movements, perhaps to make lines of parallel motion in the sand. My actions mesh with its movements, and they could do so in respect of their nature/direction/timing/intensity. Now perhaps it is not really a robot controlled by some very simply pre-programmed movements, but actually is a genuine agent. Its movements are, unbeknownst to me, genuine actions. Then Neisser’s sufficient condition for intersubjectivity is fulfilled: each of our actions does mesh with the other’s. But this is not plausibly a case of intersubjectivity, on any intuitive understanding of that notion. Intersubjectivity must at a minimum involve the recognition by each party to the interaction of the status of the other as agent, and as acting on the particular occasion in question. In interactions that are manifestations of intersubjectivity, each participant draws, no doubt tacitly, on some conception of the other’s movements as flowing from the emotions, purposes, inclinations, desires, or needs of a continuing subject who is also the agent of the actions on the particular occasion in question. Indeed, the very notion of the “appropriateness” of a response, present in Neisser’s sufficient condition, cannot plausibly be elucidated except in terms of what seems appropriate to a subject of a certain kind, a member of a certain species with certain emotions, purposes, inclinations, desires, and needs. When, by contrast with the thing mistakenly taken for a robot, an adult perceives an infant’s smile as a smile, and smiles back and waves, and the infant responds, we do have intersubjectivity. There is intersubjectivity here because there is action perceived as

<sup>10</sup> Neisser (1988, p. 43) cites Murray and Trevarthen (1985).

action, there is an adjusted response to this action, a response itself perceived as action; and so forth. The example is an instance of what Philippe Rochat describes as “the irresistible drive *to be with others, to maintain social closeness, and to control social intimacy*” (2009, p. 2). The adult’s interaction with the infant is also continued because of each party’s pleasure at the other’s pleasure in the interaction. These more primitive examples of intersubjectivity contain the materials from which are built the more sophisticated cases of interpersonal self-consciousness (and which are not always such fun).

Autistic subjects are said not to gain positive affect from rapid interaction with other persons, in the way non-autistic subjects so often do. It seems to me this lack of positive affect should be treated as consequential upon the lack of certain representational capacities. The autistic subject suffers the lack of positive affect in such interactions because of an absence of subpersonal mechanisms with rich representational contents concerning the mental states of his co-participant, including mental states with contents that involve the kinds of embeddings making reference to the first person that we mentioned earlier. When, however, the affect is present as in normal subjects, the pleasure taken in the fact of the other’s pleasure in the interaction is something that can iterate up the hierarchy. Such pleasure can itself involve instances of interpersonal self-consciousness. One can be aware, and take pleasure in the fact, that the other is taking pleasure in the fact that one has certain attitudes and occurrent emotions with first person contents. Some cases will fit Barwise’s self-involving form. We can speak of there being a situation *S* such that:

- S*: you and the other interact in such-and-such fashion
- S*: you take pleasure in the existence of *S*
- S*: the other takes pleasure in the existence of *S*.

Neisser goes on to write of “the self that is established in these interactions” (p. 43). If that means something ontological, something at the level of reference—that there is a self that does not exist independently of such social interactions—that seems hard to defend. It seems that one and the same subject who does not at an earlier time have the capacity for intersubjective interactions may later gain that capacity. That does not bring a new subject into existence. It rather enriches the capacities of a continuing subject who was there already. Perhaps the talk of selves is merely a *façon de parler*, and the apparent introduction of an ontology should not be taken too seriously.

There is, however, a closely related thesis that does not need to be formulated in metaphysical or ontological terms, a thesis endorsed by Neisser in the same passage as that just quoted, and which can be understood as concerning only modes of representation, rather than what is represented. Neisser writes that Gibson’s “principle that all perceiving involves co-perception of environment and self applies also to the *social* environment and to the *interpersonal* self....Just as the ecological self is specified by the orientation and flow of optical texture, so the interpersonal self is specified by the orientation and flow of the other individual’s expressive gestures” (p. 43). Even in the case of perception of the spatial, physical

environment, there can be subjects who enjoy such perceptions without using first person contents that place themselves in the world. These are the subjects I described in “Subjects and Consciousness”, mentioned earlier above in passing, who enjoy perceptual states with demonstrative perceptual contents concerning some of the objects and events around them, who use a *here* and a *now* but who do not use any form (even nonconceptual) of the first person. In the terminology of that paper, they are the subjects with degree 0 of involvement of any self-representation in their conception of the objective world. Even when such a subject makes a transition to representing itself as having a location in the spatial world, the notion of self-representation it then employs does not, I would argue, have to be explained in terms of spatial location. It can be explained in terms simply of a notion or concept that, *de jure*, refers to the subject in whose mental state it is featuring. (I realize that I here contradict a long Oxonian tradition, running through Peter Strawson, Gareth Evans, and John McDowell, and that their views on this need to be addressed separately.) To say, what is true, that in mature human perception, spatial content concerns both the environment and the subject, is not to say or imply that: first person representation has to be explained in part in terms of the spatial content of experience; nor is it to say or imply that being a subject is to be elucidated in terms of such spatial experience.

I suggest that structurally similar points apply to the subject who is interpersonally self-conscious, and to the first person way in which such a subject thinks about himself, even when engaged in intersubjective interactions. The first person notion, which I emphasized is both used and thought about in the characterization of ascriptive interpersonal self-consciousness, is the *same* first person notion that enters the content of non-social perception. So I am in disagreement with Neisser’s thesis, even if it is construed merely as one about concepts or notions, rather than as one about their referents. No new first person notion or concept is involved in a transition to the intersubjectivity involved in interpersonal self-consciousness. Some strange consequences would follow if a new concept or notion were involved. A subject who later acquires a sufficiently rich conception of those with whom he interacts to be interpersonally self-conscious may remember his earlier encounters with objects and events, prior to his acquiring this richer conception. It seems wholly natural to say that these earlier memories have a first person content, “I was building a snowman” for instance. Would there be some fallacy in this subject’s saying to someone, on the basis of these memories, now that he is interpersonally self-conscious, “I once built a snowman”? It is hard to see that this is anything but correct and reasonable. It does not rest on any identity inference involving two different species of first person concepts or notions. It is not as if this subject had to distinguish between two different kinds of belief and knowledge, with two different kinds of first person concept involved in each. There is only one first person concept that he employs, both before and after his acquisition of richer psychological and interpersonal conceptions. The newly acquired conceptions are richer conceptions of his relations to other subjects, and not literally newly acquired first person concepts.

These remarks do not apply only to Neisser, whose talk of various kinds of self is an instance of a tradition extending both backwards and forwards from him, in both

developmental psychology and in sociology. These remarks apply equally to talk elsewhere in that tradition of ‘the social self’, ‘the presented self’, ‘the self represented by others’, and the like. In all these cases, I would say that the important claims made using such vocabulary should be regarded not as involving a new kind of entity, nor a new kind of first person concept, but rather as claims concerning how just one subject thinks of himself (ordinary first person), or perceives himself (ordinary first person), as represented by others. No additional ontology or first person is required for the description and explanation of the phenomena.

One could of course stipulatively introduce an ontology of intentional objects, interpersonal selves that have no nature beyond what is given in the content of the interpersonal attitudes of a particular subject. Introducing or articulating such an ontology cannot settle substantive philosophical issues about whether the same subject who did not previously enjoy interpersonal self-consciousness now does so. Nor can it settle whether such a subject is using the same first person notion or concept as is deployed outside intersubjective interactions.

## 5 The relations of interpersonal self-consciousness

What is the relation of interpersonal self-consciousness to perspectival self-consciousness and to reflective self-consciousness? Does it presuppose either or both of them?

Let us start with the question of whether the interpersonal variety of self-consciousness presupposes the perspectival. Though there are certain kinds of perspectival self-consciousness associated with the mirror-neuron representation of actions, it does not seem that anything beyond that is involved in simple instances of interpersonal self-consciousness. Consider an example. I can tell that you are looking at me, and I can reach an awareness that you know that I know that I am standing on your foot. This does not seem to require that I adopt a third person perspective on myself. It seems only to involve the attribution to you of knowledge of spatial properties I perceive myself to have. This perception seems to involve only a first person perspective.

I visually perceive my foot as on top of yours. You also perceive it as standing in that relation to your foot. The way in which my foot is given in perception to me (both visually and proprioceptively) is very different from the way in which it is given to you (both in vision and in touch). But I neither need to know, nor do I at any level have to represent, the way in which you experience these two objects. My foot and yours are such that I know of them that you know I know the first to be on top of the second. If I reach this knowledge via the mechanisms of joint awareness, although there will be ways in which these objects are given to you, those ways do not need to enter the processes involved in joint awareness. The existence of ways in which you perceive these objects, and their systematic variation with the ways I perceive them, is something on which this interpersonal self-consciousness relies. But interpersonal self-consciousness does not, and does not need to, represent the systematic variation: reliance of a computational process on something is to be distinguished from representing it. I end up in a state of knowing that you perceive

of one object that it stands in a certain relation to another object without engaging in computations about the ways in which those objects are given to you. No doubt some subjects could engage in such computations, but humans do not need to in cases like these.

This relative independence of interpersonal self-consciousness from perspectival self-consciousness contributes to an explanation of how it can be the case that for a subject, the social world, and how he is represented by others in that social world, including his attitudes about himself, can be real and possibly vivid for him, even if he has a rather limited range of knowledge of those of his properties that he could only know about in a third person way. Appreciation of oneself as represented in the mental states of others, and as having attitudes, can be present in a subject even if the subject lacks the rather modest ability to gain knowledge about himself in such third person ways as looking in a mirror, or observing the shape of his shadow. The ability to represent “they are looking at me”, and “they are doing such-and-such in relation to me”, do not imply the ability to gain knowledge about oneself from a third person perspective on oneself.

There is a different relation that holds between interpersonal self-consciousness and perspectival self-consciousness. When a subject is capable of the interpersonal self-conscious states involved in linguistic communication (or indeed in non-linguistic communication), such communication can convey to the subject information that he would otherwise be able to attain, if at all, only by an exercise of perspectival self-consciousness. Information that the other person has about you is, in the nature of the case, obtained in a third person way. You are the third person, in such cases. Facts about yourself that you could know only in a third person way, that is, in a way that would equally in normal circumstances could equally give you information about someone other than yourself—such as inferring that there is a paint on your back because you have leaned against a newly painted wall—will be available from helpful interlocutors, without your having to engage in those ways of coming to know.

What of the other question, of whether interpersonal self-consciousness presupposes reflective self-consciousness? Consider first the case in which the means by which a state of interpersonal self-consciousness is attained is the operation of the second order simulation described earlier. In simple cases, this does not require any exercise of reflective self-consciousness.<sup>11</sup> Suppose it is your interpersonal self-consciousness that is in question. Then the second step of that simulation involves assessment of whether the other, given his knowledge of your situation, would judge that if he were your situation, he (the other) would think himself to be so-and-so. This must involve your making an assessment of what he would judge, given certain perceptions and background information (and possibly misinformation). That in turn seems to involve your making a first person assessment of what you would judge, given certain perceptual states and background information (and misinformation). But although I have just described the operation of this procedure in terms of what you would judge given certain perceptual states, it seems clear that you would not even

---

<sup>11</sup> My thanks to Michael Martin for setting me straight on this.

need a concept or a notion of perception to go through this procedure yourself. For you need only to think of what you would judge in circumstances described as objective states of affairs in which there are things and events of such-and-such objective kinds standing in so-and-so relations to you. Your ability to think of these things, events and relations may be dependent, constitutively, on your capacity to perceive them in various ways. But it does not follow that to think of them in these objective ways, you must also be capable of thinking about those experiences. (If that were so, a subject could not have observational concepts without also having concepts of perceptual states; but in fact a subject could do so.) You think of the objects, events and the rest under concepts made available by your capacity to perceive. That is to be distinguished from thinking about perceptions. So interpersonal self-consciousness attained by second-order simulation does not need to involve exercises of reflective self-consciousness.

There may, however, be arguments by a different route that interpersonal self-consciousness involves reflective self-consciousness. We noted that a case can be made that interpersonal self-consciousness attained by mechanisms of joint attention may in one way or another be more basic than interpersonal self-consciousness attained by second-order simulation. Joint attention characteristically involves awareness of each participant of the other participants' perceptual states, and it is plausible that there is a first person component, involving the ability to self-ascribe, in grasp of any notion or concept of perception. If all that is true, then at least for humans like us, there will be no interpersonal self-consciousness without the self-ascriptive abilities involved in reflective self-consciousness.

## 6 Concluding remarks

The literature on the philosophy of mind abounds with theses that either directly or indirectly involve interpersonal self-consciousness. These theses include the idea that the concept of a subject other than oneself who also employs the first person concept needs to be explained in terms of second person thought.<sup>12</sup> They include also theses in somewhat different traditions, such as Sartre's claim that "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" (1992, p. 361, in the section titled 'The Look'). Though I am sceptical of many such theses, the bearing of the present paper upon them is that they should all be assessed using the account of interpersonal self-consciousness that I have outlined here.

Theses about interpersonal self-consciousness are by no means exhausted by concerns about the metaphysics of the subject, or the nature of the intentional content involved in first person mental states, or the epistemology of self and others.

<sup>12</sup> Thus Rödl, *Self-Consciousness*, p. 187: "thought about another self-conscious subject is a thought whose linguistic expression requires use of a second-person pronoun". In Rödl's terminology, a self-conscious subject is simply a subject capable of first person thought. Rödl uses the terminology mentioned at the start of this paper.

Interpersonal self-consciousness is also foundational in our emotional lives and in what we value. Martin Buber (1958, p. 17) writes “I become through my relation to the *Thou*; as I become *I*, I say *Thou*. All real living is meeting”. One can find something in his points, even if one is highly sceptical if such sentences are taken as theses in the metaphysics of the conscious subject, or as theses in the theory of intentional content. Buber’s claims can instead be taken as points about what matters to us, about some of our values and our emotional lives. We must also draw on an account of interpersonal self-consciousness in that still relatively unexplored territory.

**Acknowledgments** This paper originated in one of my summer seminars at University College London; it was further developed for one of my Kohut Lectures at the University of Chicago early in 2011, and for one of David Chalmers’ Consciousness Project Meetings at NYU in the fall of 2011. I thank Jonathan Lear for discussion in Chicago, and Ned Block and David Chalmers, and other members of the discussion group, for comments at NYU. A version was also presented to the Oberlin Philosophy Colloquium in May 2012, the fortieth in that now famous series. At Oberlin I learned from the exemplary, insightful prepared commentary of Michael Martin; I note in the text one correction I have made in the light of his comments. I thank also Fred Dretske for illuminating discussion at the same occasion. A more recent version of this material was presented at the Workshop on the Self, organized by Natalie Sebanz and Hong Yu Wong at the Central European University in Budapest in the summer of 2012. That occasion too resulted in revisions. I am particularly grateful for conversations in Budapest with the developmental psychologist Philippe Rochat.

## Appendix

Suppose we have a case in which you are interpersonally self-conscious with respect to me, CP. In such an example, you are aware that CP attributes to you a first person thought, say that you are F. I continue to use the neo-Fregean notation of note 5 above. The first person thought you would have to have for this attribution to be correct is one you could express by saying “I’m F”. In the neo-Fregean notation, an attribution of this content to your thinking would be regimented:

(A) Think (you, [self]<sub>you</sub><sup>^</sup><is F>).

When you are interpersonally self-conscious with respect to me, CP, you are aware that I (CP) am attributing thinking specified in (A) to you. For this to be true, it has to be that:

(B) Thinks (CP, <Think><sup>^[that person]<sub>NN</sub><sup>^</sup>[[self]]<sub>that person</sub><sup>^</sup><<is F>>).</sup>

Here we have the double embedding shown by the double square brackets, the canonical sense of the first person sense, [self]<sub>NN</sub>, that is employed by you. Finally, since you are aware that (B) holds, we have the triple embedding as correct: in neo-Fregean notation,

(C) Aware-that (you, <Thinks><sup>^[that person]<sub>CP</sub><sup>^</sup><<Thinks>><sup>^[self]<sub>NN</sub><sup>^</sup>[[self]]<sub>NN</sub><sup>^</sup><<<is F>>>).</sup></sup>

Enthusiasts for this sort of issue will realize that I have here made the occurrence of ‘[self]<sub>NN</sub>’ transparent, which it is in such attributions. From your being aware that I

think that you think that you yourself are F, it does follow that: you are aware of someone  $x$  (viz. yourself) of whom you are aware that I think of that person  $x$  that he thinks himself to be F.)

All this is involved in the formal representation of the fact that in interpersonal self-consciousness, you are aware of me as treating you as self-representing.

## References

- Abbott, J. (1981). *In the belly of the beast: Letters from prison*. New York: Random House.
- Barwise, J. (1988). Three views of common knowledge. In M. Vardi (Ed.), *Proceedings of the second conference on theoretical aspects of reasoning about knowledge* (pp. 365–379). Los Altos, CA: Morgan Kaufmann.
- Barwise, J. (1989). *The situation in logic (CSLI lecture notes 17)*. Stanford, CA: CSLI.
- Buber, M. (1958). *I and Thou*, tr. Ronald Gregor Smith. London: Continuum Press.
- Fodor, J., & Pylyshyn, Z. (1981). How direct is visual perception? Some reflections on Gibson's "ecological approach". *Cognition*, 9, 139–196.
- Gallese, V. (2005). "Being like me": self-other identity, mirror neurons, and empathy. In S. Hurley & N. Chater (Eds.), *Perspectives on imitation: From neuroscience to social science (Vol. 1)*. Cambridge, MA: MIT Press.
- Gallup, G., Jr. (1970). Chimpanzees: Self-recognition. *Science*, 167(3914), 86–87.
- Goffman, E. (1959). *The presentation of self in everyday life*. New York: Doubleday.
- Grandin, T. (2006). *Thinking in pictures: My life with autism*. New York: Random House.
- Heck, R. (2002). Do demonstrative have senses? *Philosopher's Imprint*, 2, 1–33.
- James, W. (1918). *The principles of psychology*. New York: Holt.
- Lewis, D. (1969). *Convention*. Cambridge, MA: Harvard University Press.
- Mead, G. H. (1934). *Mind, self, and society*. Chicago: Chicago University Press.
- Murray, L., & Trevarthen, C. (1985). Emotional regulation of interactions between two-month-olds and their mothers. In T. Field & N. Fox (Eds.), *Social perception in infants*. Norwood, NJ: Ablex.
- Neisser, N. (1988). Five kinds of self-knowledge. *Philosophical Psychology*, 1, 35–59.
- O'Brien, L. (2012). The ordinary notion of self-consciousness. In J. Liu & J. Perry (Eds.), *Consciousness and the self: New essays*. Cambridge: Cambridge University Press.
- Peacocke, C. (1981). Demonstrative thought and psychological explanation. *Synthese*, 49, 187–217.
- Peacocke, C. (2005). Joint attention: Its nature, reflexivity, and relation to common knowledge. In N. Eilan, C. Hoerl, T. McCormack, & J. Roessler (Eds.), *Joint attention: Communication and other minds*. Oxford: Oxford University Press.
- Peacocke, C. (2010). Self-consciousness. *Revue de Métaphysique et de Morale Issue*, 4, 521–551.
- Peacocke, C. (2012). Subjects and consciousness. In A. Coliva (Ed.), *The self and self-knowledge*. Oxford: Oxford University Press.
- Perry, J. (1979). The problem of the essential indexical. *Noûs*, 13, 3–21.
- Reiss, D., & Marino, L. (2001). Mirror self-recognition in the bottlenose dolphin: A case of cognitive convergence. *Proceedings of the National Academy of Sciences*, 98(10), 5937–5942.
- Rochat, P. (2009). *Others in mind: Social origins of self-consciousness*. New York: Cambridge University Press.
- Rödl, S. (2007). *Self-consciousness*. Cambridge, MA: Harvard University Press.
- Sartre, J.-P. (1992). *Being and nothingness tr. H. Barnes*. New York: Simon and Schuster.
- Schiffer, S. (1972). *Meaning*. Oxford: Oxford University Press.