

# **Self-Consciousness and Embodied Experience**

Thesis submitted to the Department of Philosophy  
for the degree of Doctor of Philosophy

Joel Alexander Smith

Department of Philosophy  
University College London  
Gower Street  
London  
WC1E 6BT

October 2003

## Abstract

The Body Claim states that a transcendental condition of self-consciousness is that one experience oneself as embodied. The contention of this thesis is that popular arguments in support of the Body Claim are unconvincing. Understanding the Body Claim requires us to have a clear understanding of both self-consciousness and embodied experience. In the first chapter I lay out two different conceptions of self-consciousness, arguing that the proponent of the Body Claim should think of self-consciousness as first-person thought. I point out that since arguments for the Body Claim tend to proceed by stating putative transcendental conditions on *self-reference*, the proponent of the Body Claim must maintain that there is a conceptual connection between self-consciousness and self-reference. In the second chapter I argue against views, originating from Wittgenstein and Anscombe, which reject this connection between self-consciousness and self-reference. In chapter three I show that a well known principle governing the ascription of content, that which Evans calls ‘Russell’s Principle’, occupies an ambiguous position with regards to the Body Claim. I argue that Russell’s Principle should be rejected. Chapter four distinguishes between two conceptions of embodied experience: bodily-awareness and bodily self-awareness. I argue that there is no such thing as bodily self-awareness and so it cannot be a transcendental condition of self-consciousness. Chapter five looks at, and finds wanting, arguments for the Body Claim that can be found in the work of Strawson. Chapter six argues that it is a transcendental condition of self-consciousness that one enjoy spatial experience. Chapters seven and eight assess two influential arguments that attempt to complete a defence of the Body Claim: the solidity argument and the action argument. I argue that neither argument is convincing. Although the conclusions are primarily negative, much is learned along the way about the nature of both self-consciousness and embodied experience.

# Contents

<i>Acknowledgements</i>	5
0. Introduction	6
1. Self-Consciousness	
1.1 Self-Awareness	11
1.2 First-Person Thought	13
1.3 Self-Reference	15
2. No Reference Views	
2.1 First-Person Guarantees	20
2.2 Guaranteed Reference and First-Person Authority	21
2.3 Immunity to Error Through Misidentification	23
2.4 Another Brand of Immunity to Error Through Misidentification?	26
2.5 The No Reference View	29
2.6 Non-Factualism	31
2.7 The Frege-Geach Problem	39
2.8 Fictionalism	42
2.9 Explaining the First-Person Guarantees	45
3. Russell's Principle	
3.1 The Ambiguous Status of Russell's Principle	49
3.2 What Russell's Principle Is	50
3.3 Evans on Singular Thought	51
3.4 'I'-Ideas	55
3.5 The Possibility of Massive Reference Failure	62
3.6 Against Russell's Principle	72
4. Embodied Experience	
4.1 Bodily-Awareness	79
4.2 Bodily Self-Awareness	82
4.3 Cassam's Identity Argument	84
4.4 Bodily-Awareness and Self-Awareness	89
4.5 Bodily Sensations	93
4.6 Experience and Belief	94
4.7 Interlude	96
5. Strawson	
5.1 The Objectivity Condition	98
5.2 Sketch of an Over Hasty Argument	102
5.3 Strawson's Objectivity Argument	102
5.4 Empirically Applicable Criteria of Identity	109
5.5 Bennett's Realism Argument	112
5.6 Strawson Against Dualism	114

6. Self-Location	
6.1 Thinking of Oneself as a Particular	120
6.2 Thinking of Oneself as Spatially Located	122
6.3 Spatial Concepts and Spatial Experience	125
7. Solidity	
7.1 Spatial Experience and Physical Objects	129
7.2 Solidity and Force	133
7.3 Touch and Bodily-Awareness	139
7.4 Objections	145
8. Action	
8.1 Perception and Action	148
8.2 Frames of Reference	152
8.3 Viewer Centred Frames of Reference	158
8.4 Egocentric Frames of Reference	161
8.5 Having a Meaning for the Subject	164
8.6 Spatial Action and Bodily-Awareness	166
8.7 Having a Body	170
9. Conclusions	175
10. Appendix	177
<i>References</i>	181

## Acknowledgements

I would like to thank the generosity of the Arts and Humanities Research Board for providing me with financial assistance throughout my research on this project.

Thinking about self-consciousness is enormously hard and I have benefited a great deal from discussions with numerous members of the Philosophy Department at University College London. In particular I must thank my supervisor, Lucy O'Brien, for all the thoughtful comments, clear guidance and general good philosophical sense she has presented me with over the past three years. I owe too a great debt to Mike Martin, who has always been incredibly supportive. I would also like to thank José Zalabardo and Mark Kalderon who have both helped me to think more clearly about many issues. Papers containing material from this thesis were read at conferences in London, Dubrovnik, Glasgow and Bristol. I am grateful to the audiences at those conferences for many helpful comments.

Since I always save the best till last, I would like to thank Ann Whittle who has read through more rough drafts and listened to me mutter incomprehensibly about Russell's Principle more times than I care to remember. Without her I would probably have given it all up years ago.

## 0. Introduction

Mature human beings are self-conscious. This is beyond doubt. Our ability to think of ourselves *as* ourselves is, say some, what separates us from mere beasts. But if self-consciousness is an aspect of our experience which sets us apart from other animals, embodied experience is one which brings us together. Both human beings and other animals experience themselves as embodied. Human beings have bodies and we experience ourselves as such.

Embodied experience is to be distinguished from embodiment. In saying that human beings are embodied, I am saying that human beings have bodies. This is a claim about the way the world is. In saying that human beings have embodied experience, I am saying that human beings experience themselves as embodied. This is claim about the way the world appears to us. To see the distinction, consider the brain-in-a-vat fantasy, in which a subject enjoys embodied experience without being embodied. Conversely, assuming that they lack experience, slugs are embodied without enjoying embodied experience.

What is the relationship between self-consciousness and embodied experience? An appealing answer to this question is that embodied experience is a transcendental condition of self-consciousness. That is, a self-conscious subject must enjoy embodied experience. Or, in other words, it is not possible that a self-conscious subject fail to enjoy embodied experience. I shall refer to this claim as the Body Claim. To repeat, the Body Claim does not state that a self-conscious subject must be embodied, but rather than a self-conscious subject must experience itself as embodied.

This thesis is a critical examination of the Body Claim. I assess arguments that have been put forward in support of that claim. But my aim is limited. I look at only those arguments which proceed via the notion of self-reference. That is, it is claimed by some that self-reference is a transcendental condition of self-consciousness, and that embodied experience is a transcendental condition of self-reference. I claim that the various arguments of this form that have been put forward all fail. Whilst it is true that self-consciousness necessarily involves self-reference, it has not been shown that embodied experience is a transcendental condition of self-reference. Therefore, the Body Claim has not been established.

I do not claim that the Body Claim is false. My position with regards to the Body Claim is agnostic. Neither do I claim that there are no successful arguments that proceed in a more direct fashion, bypassing the notion of self-reference (although I have yet to see one). It may seem that the scope of the thesis is, as a result, overly narrow. I hope that this is not the case. First, the arguments that I will be discussing have been enormously influential and are by far the most popular ways of defending the Body Claim. Second, the arguments are broad, in the sense that they range through a great many topics within the philosophy of mind and language.

The Body Claim is a transcendental claim. Transcendental claims are naturally characterised as claims to the effect that  $x$  is a necessary condition of  $y$ , where  $y$  is a feature of our experience that is beyond sceptical doubt. The Body Claim fits the bill nicely: self-consciousness is a feature of our experience the existence of which is beyond sceptical doubt, and embodied experience is claimed to be a necessary condition of it. But as natural as this characterisation is, there are reasons to think that it should be amended. The Body Claim denies the possibility of a self-conscious subject without embodied experience. But the methodology employed in traditional arguments for the Body Claim involve conceivability. It is argued that it is not conceivable that a self-conscious subject lack embodied experience. The legitimacy of this clearly relies on the existence of a strong link between conceivability and possibility. Specifically, it relies on the principle which states that that which is not conceivable is not possible. And this principle is highly controversial<sup>1</sup>. We do not want the Body Claim, and transcendental claims in general, to be hostage to the fortunes of this principle.

A better way to characterise transcendental claims is explicitly as claims of conceivability. Transcendental claims should not be thought of as of the form ' $x$  is a necessary condition of  $y$ ', but rather as of the form ' $y$  without  $x$  is inconceivable'. Discovering that  $x$  and  $y$  are linked in this way is still philosophically interesting, for it tells us something maybe not about modal reality, but about the relations between our concepts; about our conceptual scheme. This is, I suggest, in line with the way in which proponents of transcendental claims have often seen their own activity.

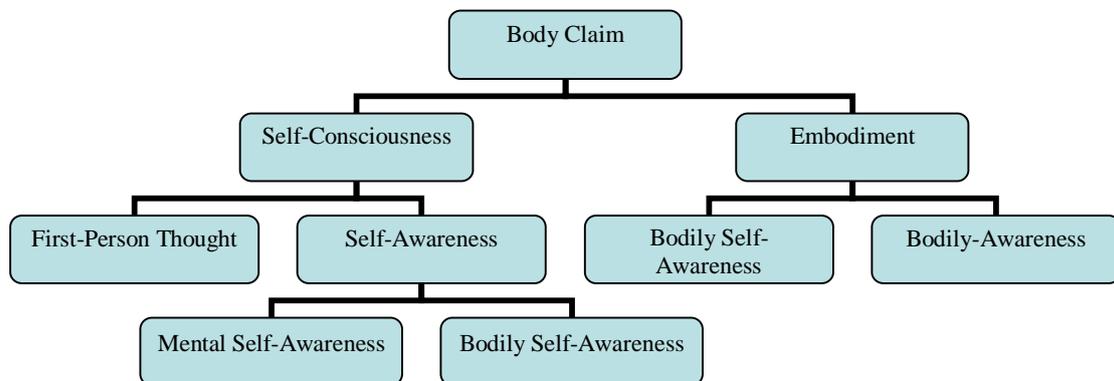
---

<sup>1</sup> Particularly so since Kripke, 1980. These issues are discussed in fascinating detail in Gendler and Hawthorne, 2002.

Transcendental claims do not claim to describe the world but “to describe the actual structure of our thought about the world” (Strawson, 1959, p.9).

We should, then, reformulate the Body Claim as the claim that it is inconceivable that a self-conscious subject lack embodied experience. This falls short of making an explicit claim about modal reality. Such a claim would need supplementation by a principle linking conceivability and possibility. I shall mark this point by speaking not of necessary conditions but of transcendental conditions. That is, as I use it,  $x$  is a transcendental condition of  $y$  if and only if  $y$  is not conceivable without  $x$ . Understood in this way, the Body Claim maintains that embodied experience is a transcendental condition of self-consciousness.

The two central concepts with which this thesis is concerned are self-consciousness and embodied experience. Getting clear about these notions is vital to any examination of the Body Claim. Throughout the following chapters a picture will emerge as to the different ways both self-consciousness and embodied experience might be understood. Roughly speaking, there are two ways to understand self-consciousness. The first is that self-consciousness is an introspective awareness of the self. This might be an awareness of the self as either a mental or a bodily thing. On the other hand, self-consciousness can be understood as a property of certain thoughts. In a similar way, embodied experience can be understood in two ways. The first is simply as the awareness that each of us has of his or her own body; bodily-awareness. The second is as a form of bodily introspection, or self-awareness. This can be represented as follows:



*fig.1*

The first thing to notice concerning this representation is that the phrase ‘bodily self-awareness’ occurs twice. This is no mistake. As will be seen, those who believe in bodily self-awareness think of it both as a form of self-consciousness and as a form of embodied experience.

Looking ahead, in the first chapter I distinguish the different ways in which self-consciousness can be understood. I begin by saying more about the distinction between self-awareness and first-person thought. I argue that any interesting version of the Body Claim must look to the second of these construals of self-consciousness. I then move on to the relationship between self-consciousness, understood as first-person thought, and the notion of self-reference. I distinguish between different strengths of the commitment to self-reference and explain what must be maintained by anyone who wishes to argue for the Body Claim via the notion of self-reference. I shall sometimes refer simply to defenders of the Body Claim, this should be taken to be restricted to defenders of the Body Claim via the notion of self-reference.

Chapter two offers a critique of the ‘no reference view’ of first-person thought. Versions of this view have been offered by both Wittgenstein and Anscombe. Associated with this position is the idea that the self-ascription of mental states should be given a non-factualist reading. I explain why the defender of the Body Claim via self-reference must reject the no reference view and show how this can be done.

The third chapter is, in many ways, the heart of the thesis as a whole. It concerns, what has come to be known as Russell’s Principle. I show that Russell’s Principle is of fundamental importance in some influential arguments for the Body Claim. However, Russell’s Principle occupies an ambiguous and unstable position. For if Russell’s Principle is accepted it poses a serious threat to the tight link between self-consciousness and self-reference that is required for the defence of the Body Claim. I argue for this by way of a critical exposition of Evans’ theory of first-person thought. The way out of the morass is the rejection of Russell’s Principle. With Russell’s Principle out of the way, the defender of the Body Claim is free to assert the strong conceptual connection between self-consciousness and self-reference.

Chapter four moves from self-consciousness to embodied experience. The two conceptions of embodied experience are distinguished and described. I then discuss the defence of the Body Claim recently offered by Cassam. Cassam understands embodied experience in the strong sense, as bodily self-awareness. First I show that his argument is flawed. Second I show that there is no such thing as bodily self-

awareness. If the Body Claim is defensible, embodied experience must be understood weakly, as bodily-awareness.

Chapter five looks for an argument for the Body Claim in the work of Strawson. I argue that none of Strawson's arguments are convincing. However, although Strawson's arguments are unsuccessful, they point towards a better argument. I dub this the self-location argument, and in chapter six I defend it. The self-location argument shows that a self-conscious subject must experience him or herself as spatially located. This is an achievement, but it does not yet vindicate the Body Claim. For it may be suggested that a self-conscious subject could experience him or herself as a spatially located but bodiless 'point of view'.

The final two chapters consider, what I take to be, the most serious arguments against the possibility of such a point of view. The first argument, discussed in chapter seven, relies on an account of the notion of solidity and the role it plays in our experience of the world as spatial. The second argument, the topic of chapter eight, offers an account of spatial action and its relation to both bodily-awareness and our ability to give content to the idea that perception is spatial. Unfortunately for the proponent of the Body Claim, neither of these arguments are successful.

My conclusion is that it has not been shown that self-consciousness without embodied experience is inconceivable. But this is not to deny that embodied experience plays a fundamental role in determining the way in which we experience the world. Even if the link between self-consciousness and embodied experience is contingent, it is plausible to think that it is, if I may be allowed this phrase, 'a deep contingency'. Furthermore, throughout the discussion of the various attempts to establish the Body Claim we learn much about both self-consciousness and embodied experience.

# 1. Self-Consciousness

## 1.1 Self-Awareness

I distinguished above between two ways of understanding self-consciousness. The first is as an introspective awareness of the self. The second is as first-person thought. I say more about first-person thought in the following sections. In this section I discuss self-consciousness understood as an introspective awareness of the self, for which I reserve the term ‘self-awareness’.

Self-awareness is an introspective awareness of the self as a thing. The hypothesis that we are self-aware is the hypothesis that when one introspects one is directly aware of something which is presented to one as oneself. There are at least two ways in which philosophers have understood this view. The first is that, in introspection, one is aware of oneself as a purely mental thing, as a bearer of mental properties only. The second is that, in introspection, one is aware of oneself as a (partly) physical thing, as a bearer of (at least some) physical properties. I will argue that understanding self-consciousness as mental self-awareness deprives the Body Claim of much of its plausibility, whilst thinking of self-consciousness as bodily self-awareness trivialises it. The proponent of the Body Claim would do well to move to an understanding of self-consciousness as first-person thought.

Since Hume it has been common to deny that there is such a thing as mental self-awareness. Hume, of course, famously wrote that,

For my part, when I enter most intimately into what I call *myself*, I always stumble on some particular perception or other, of heat or cold, light or shade, love or hatred, pain or pleasure. I never can catch *myself* at any time without a perception, and never can observe anything but the perception. (Hume, 1978, pp.252)

There are a variety of interpretational issues to be settled concerning this passage<sup>1</sup>, but on the face of it Hume appears to be denying what the proponent of self-awareness is asserting. Namely that introspection reveals an awareness of the self as an object. Presumably, furthermore, Hume is concerned to deny that we are introspectively

---

<sup>1</sup> Nice discussions of which can be found in Craig, 1987, Ch.2, Penelhum, 2000, and Pitson, 2002.

aware of ourselves as mental things, the idea that we might be introspectively aware of ourselves as physical things was surely not his target.

Philosophers have long been divided as to whether Hume was correct to reject mental self-awareness<sup>2</sup>. I tend to side with Hume. A simple line of reasoning that can move one to this position begins with the thought that the only reason to think that there might be an introspective awareness of the self as a mental object would be if it could explain the rather special knowledge that we have of our mental properties. As Shoemaker has written,

Presumably it will be pointless, at best, to suppose that there is introspective perception of the self unless this perception plays some role in explaining our introspective self-knowledge – our knowledge of our own mental states. The most straightforward account would be this: I know that I have thus and such a mental state – that I am angry, in pain, or desirous of a drink – because I introspectively observe myself having it. (Shoemaker, 1986, p.13)

The problem, however, is that it does not seem especially plausible to say that the way in which we come to know that we have thus and such mental state is by an introspective awareness of the self as being thus and such. As Evans so eloquently put it, “in making a self-ascription of belief, one’s eyes are, so to speak, or occasionally literally, directed outward – upon the world.” (Evans, 1982, 225). The model of self-knowledge according to which I know what I believe by an ‘inward glance’ at myself, seems at odds with our actual practice according to which I know what I believe by observing the world around me.

Of course, simply making this observation is not to present an argument against the mental self-awareness view. The proponent of that view will immediately point out that, whilst it may be inappropriate to treat knowledge of our beliefs on the ‘inward glance’ model, knowledge of our desires, hopes, fears, emotions, and sensations is another matter. This may or may not be the case. But as long as the proponent of mental self-awareness is willing to grant Evans’ point about beliefs the damage is, to a certain extent, done. For, if this concession is made, at least some self-knowledge is not to be explained via an introspective awareness of the self. This

---

<sup>2</sup> Concurring with Hume are both Kant, 1997, A107 & B408, and Schopenhauer, as quoted in Janaway, 1989, p.120. More recently the mental self-awareness view has been rejected by Shoemaker, 1986 & 1994, and Martin, 1997. Those who appear to reject the Humean line include Russell, 1910-11, (although by the time of Russell, 1921, p.141, he has clearly changed his mind), Chisholm, 1969, and G. Strawson, 1997.

poses a threat to the general explanatory utility of mental self-awareness, for if it is not needed to explain some self-knowledge why should we think it is needed to explain the rest?

It seems that we have good reason to want to dissociate the Body Claim from the understanding of self-consciousness as mental self-awareness. The very fact that so many philosophers have followed Hume in denying that we have such an awareness is enough to give us pause. And if it turns out, as I have been suggesting, that mental self-awareness has little or no explanatory role to play in accounting for self-knowledge, we may have good reason to be of the same mind as Hume. But, of course, if there is no mental self-awareness there is little interest in giving its transcendental conditions.

There is a far more pressing reason to dissociate the Body Claim from an understanding of self-consciousness as bodily self-awareness. The Body Claim states that embodied experience is a transcendental condition of self-consciousness. If we understand self-consciousness as bodily self-awareness this claim becomes the triviality that embodied experience is a transcendental condition of an introspective awareness of oneself as a physical thing. I have far more to say about bodily self-awareness in chapter four (recall that ‘bodily self-awareness’ appears twice in *fig.1*), for now I need simply point out that an interpretation of the Body Claim that does not trivialise it is to be preferred.

## **1.2 First-Person Thought**

The second way of understanding self-consciousness is as first-person thought. Understood in this way, self-consciousness is a property of certain thoughts. Self-conscious thoughts are thoughts about oneself, but not all thoughts about oneself are self-conscious. Suppose I believe that Smith is about to be attacked by a tiger. This is a belief about myself, but it is not self-conscious. It can be that although I believe that Smith is about to be attacked by a tiger, and I am Smith, I fail to believe that I am about to be attacked by a tiger, for the reason that I fail to believe that I am Smith. Self-conscious thoughts are thoughts about oneself concerning which it is not possible to fail to realise that they are about oneself. Self-consciousness is that property which accounts for this feature of some of the thoughts that one has about oneself. Roughly speaking, a self-conscious thought is one which one would express in language by

using the first-person pronoun, 'I'. For this reason, we can call such thoughts 'I'-thoughts, and this second conception of self-consciousness *first-person thought*. I will often refer to 'I'-thoughts as self-ascriptions. This is for the reason that 'I'-thoughts involve the ascription of predicates to oneself, for instance 'I have a headache', 'I am six feet tall' etc.<sup>3</sup>

Self-conscious thoughts can also be illuminatingly characterised by their functional role. It is a familiar point that self-conscious thoughts directly dispose one to act, whereas unself-conscious thoughts do not. If I believe that I am about to be attacked by a tiger, I will be disposed to cower in fear. If I believe that Smith is about to be attacked by a tiger, I will only be disposed to cower if I also believe that I am Smith<sup>4</sup>. In addition to directly disposing one to act, self-conscious thoughts are thoughts which are liable to be affected by the 'various special ways' one has of gaining knowledge about oneself<sup>5</sup>. Self-conscious thoughts are sensitive to information gained in various special ways. These special ways certainly include introspection and arguably also include memory, bodily-awareness and self-locating abilities. If I am introspectively aware of a headache, my self-conscious thoughts will adjust to accommodate this; I will come to believe that I have a headache. On the other hand, I will not necessarily come to believe that Smith has a headache, since I may fail to believe that I am Smith. I will return to this feature of self-conscious thought in chapter two.

I do not mean to suggest by these remarks that 'I'-thoughts can be analysed purely in terms of their functional role. Rather, considerations of this kind allow us to locate the phenomenon of self-conscious thought. First-person thoughts are thoughts which tend to be expressed by the word 'I', directly dispose one to action, and are sensitive to certain special ways of gaining knowledge about oneself.

It is beyond sceptical doubt that we are self-conscious in the sense that we have the capacity to think first-personally. This is good news for the defender of the Body Claim, for the Body Claim can be thought of as offering a transcendental condition of a capacity that is uncontroversially possessed by the majority of mature human beings. Another good reason for taking the Body Claim to be stating a

---

<sup>3</sup> I shall speak of mental self-ascriptions (self-ascriptions of mental predicates) and physical or bodily self-ascriptions (self-ascriptions of physical or bodily predicates).

<sup>4</sup> The link between self-conscious belief and dispositions to act is articulated in Perry, 1977 & 1979.

<sup>5</sup> This aspect of self-conscious thought is stressed in Evans, 1982, Ch.7.

transcendental condition of first-person thought is that it has more philosophical interest than has self-awareness. As Shoemaker suggests, the primary reason for postulating self-awareness is as an explanation of certain features of some first-person thoughts. First-person thought is the interesting phenomenon, and it is that to which we should primarily be paying attention. Finally, taking the Body Claim as a claim about the transcendental conditions of first-person thought helps us to understand why defenders of the Body Claim have tended to concentrate on arguments which proceed via the notion of self-reference. Self-awareness has nothing obviously to do with self-reference, unless it is by way of first-person thought. For first-person thought, it is frequently claimed, is self-referential<sup>6</sup>.

### 1.3 Self-Reference

The defender of the Body Claim is committed to there being a conceptual connection between self-consciousness and self-reference. To argue for the Body Claim by showing that embodied experience is a transcendental condition of self-reference, one must claim that the capacity to self-refer is a transcendental condition of self-consciousness. If the capacity for first-person thought does not entail the capacity for self-reference then no matter how successful one is at spelling out the transcendental conditions of self-reference, one will not have reached so far as self-consciousness.

One position that could be taken with regards to the connection between self-consciousness and self-reference is that every first-person thought is self-referential. This view can be thought of as a commitment to what I shall be calling the self-reference rule. The self-reference rule states that any first-personal thought refers to the person who produced it<sup>7</sup>. Holding this view would secure the connection between self-consciousness and self-reference that is required for the Body Claim. On the other extreme, it might be held that no first-person thought is self-referential. This 'no reference view' is the subject of the following chapter. If correct, this view would clearly undermine arguments for the Body Claim that proceed via self-reference.

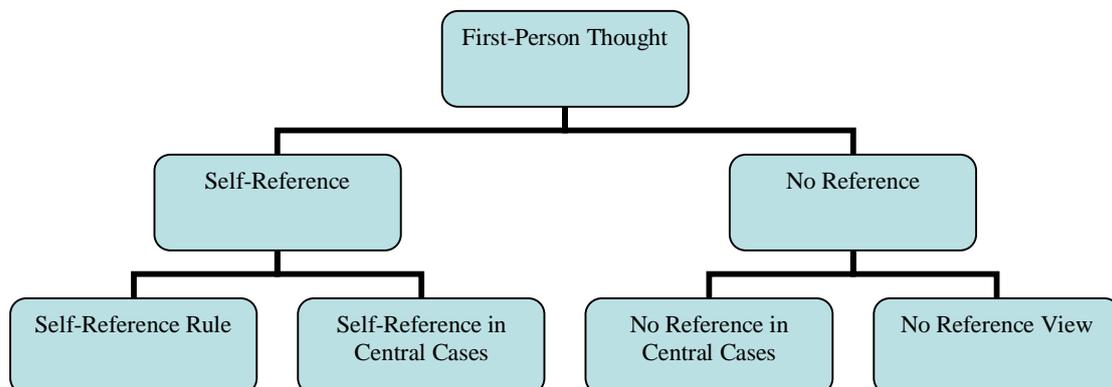
---

<sup>6</sup> Bach, 1987, pp.39-41, objects to saying that thoughts refer, claiming that reference is a four-place relation between speaker, word, audience and object. But one can linguistically refer whilst on one's own. Linguistic reference here can be thought of as a three-place relation between speaker, word and object. Similarly, reference in thought can be a three place relation between thinker, thought and object.

<sup>7</sup> The self-reference rule is the analogue in thought of what Kaplan, 1977, takes to be the character of the word 'I'. The self-reference rule is defended in Campbell, 1994a.

The debate between the defender of the self-reference rule and the proponent of the no-reference view is a debate between two extremes. There are at least two intermediate views which are relevant to the defender of the Body Claim. The Body Claim can be argued for without reliance on the exceptionless self-reference rule. For, as long as it is maintained that central cases of first-person thought are self-referential, we will still have the required conceptual connection between self-consciousness and self-reference. Similarly, it might be held that even though many first-person thoughts are self-referential, the central cases of first-person thought are not. This position would cause problems for the Body Claim if it were held that self-consciousness could be accounted for purely in terms of the non-referential central cases. For then it could be that a self-conscious subject may fail to satisfy the transcendental conditions of self-reference.

Of course something must be said about what the central cases might be. At this point I think the best way to proceed is intuitively. Intuitively, the central cases of first-person thought are self-ascriptions of mental predicates, for example the self-ascription of beliefs, desires, sensations, and experiences. It is self-ascriptions of mental predicates that have traditionally been thought to be particularly expressive of self-consciousness. Indeed, the proponent of the Body Claim must have mental self-ascriptions in mind, since the claim that the capacity to self-ascribe physical predicates requires embodied experience is far less interesting than its mental counterpart. The four positions I have been sketching can be represented as follows:



*fig.2*

Whilst I believe that there is a strong case to be made for the self-reference rule, it should be clear why holding such a strong position is not required of the defender of the Body Claim. For, the defence of the Body Claim maintains that embodied experience is a transcendental condition of self-reference. If it is further claimed that central cases of self-consciousness are self-referential it follows that central cases of self-consciousness have, as a transcendental condition, embodied experience. If it is further claimed that self-consciousness necessarily involves these central cases, the Body Claim is safe and dry. I take it, also, that the positions on the ‘no reference’ side of *fig.2* are not the intuitive views, that some form of commitment to self-reference is the default position. This means that arguing against the two forms of the no reference view is all that is required for a defence of the connection between first-person thought and self-reference. No additional argument is required.

The fact that the proponent of the Body Claim can get by without a commitment to the self-reference rule allows him or her to avoid certain tricky cases. For, there are views upon which although central cases of first-person thought are self-referential, certain atypical cases serve as counterexamples to the exceptionless self-reference. The self-reference rule faces threats from substance dualism, the nature of the imagination, and certain recalcitrant linguistic phenomena. I say a little about these here in order to clear the way for the real challenge posed by the no reference view and its ilk.

Certain views concerning the metaphysics of the self, of which substance dualism is the most obvious (but there may be others), appear to entail the falsity of the self-reference rule. According to substance dualism what I am is an immaterial soul ‘connected’ to a body. This view has no problem dealing with mental self-ascriptions which uncontroversially ascribe a mental predicate to the immaterial soul. But the dualist faces a trilemma when he or she comes to give an account of physical self-ascriptions such as ‘I am six feet tall’. For the soul is not supposed to have any non-mental properties. The three horns of the trilemma are these: (1) Claim that all non-mental self-ascriptions are false, (2) Reject the claim that what (you and) I am is an immaterial soul, in favour of the claim that the self is a ‘compound of two substances’, a body and a soul, and (3) Claim that the occurrence of ‘I’ in non-mental self-ascriptions is misleading, that it really refers to something other than oneself, i.e. one’s body.

None of these options are particularly attractive. The claim that all non-mental self-ascriptions are false is highly counterintuitive. The version of dualism which treats the self as a compound of two substances is problematic, not least in implying that what one is (the compound) is non-identical to the thing that instantiates all one's mental properties (the soul)<sup>8</sup>. This leaves the view that non-mental self-ascriptions really refer to something other than oneself, and so are not really *self*-ascriptions. And this view contradicts the self-reference rule. This view faces serious difficulties in the face of our linguistic practices<sup>9</sup>, but its refutation requires a refutation of the metaphysics upon which it is based.

Fortunately the defender of the Body Claim need not set about refuting substance dualism. For the dualist allows that the central cases of first-person thought are self-referential. Although, for various reasons, it seems unlikely that the dualist would want to defend the Body Claim, there is no reason why a dualist could not do so via the notion of self-reference.

Thinking about the imagination has led some philosophers to deny the self-reference rule. We can see why one might be tempted by this view by considering the following piece of reasoning: (1) I can only imagine that which is possible<sup>10</sup>, (2) I can imagine that I am Napoleon, (3) It is not possible that I am Napoleon. These three claims appear to be inconsistent. One tempting way of resolving the inconsistency is by claiming that the second occurrence of 'I' in the phrase 'I am imagining that I am Napoleon' does not in fact refer to oneself. Rather, one is thinking about Napoleon in the 'imaginary first-person'<sup>11</sup>. This, it would seem, contradicts the self-reference rule.

I will not comment on the strength of this line of reasoning, for there is no reason for the defender of the Body Claim to take a stance on this issue. For occurrences of 'I' that occur within the scope of 'I am imagining that...' are clearly not central cases of first-person thought. It is a fascinating question how first-person reference operates in the context of the imagination, but it is not one with which I shall be concerned.

---

<sup>8</sup> See Olson, 2001.

<sup>9</sup> See Stanley, 1998.

<sup>10</sup> See the introduction. This view is most famously held by Hume, "'Tis an established maxim in metaphysics, *That whatever the mind clearly conceives includes the idea of possible existence*, or in other words, *that nothing we imagine is absolutely impossible*." (Hume, 1978, p.32).

<sup>11</sup> See Velleman, 1996.

Finally, some philosophers have claimed that certain linguistic phenomena have shown that the word ‘I’ is not governed by a linguistic analogue of the self-reference rule<sup>12</sup>. This, coupled with some version of the thesis that thought mirrors language may be thought to pose a problem for the self-reference rule in the domain of thought. Again, this debate need not trouble the proponent of the Body Claim. For the linguistic phenomena cited are clearly all atypical uses of ‘I’, for instance leaving a note on someone else’s office door reading ‘I am not here today’. Central cases are left untouched. If the Body Claim can be defended without reliance on the self-reference rule, these cases should not be troubling. In the following chapter I shall be arguing against both the strict no reference view and the view that the central cases of first-person thought are non-referential. I claim that there is a tight conceptual connection between self-consciousness and self-reference. I leave it as an open question whether the self-reference rule is true.

---

<sup>12</sup> See, for example, Corazza, Fish and Gorvett, 2002.

## 2. No Reference Views

### 2.1 First-Person Guarantees

It has frequently been noted that the word ‘I’, and corresponding ‘I’-thoughts, are subject to guarantees against a number of kinds of error. It has been argued these guarantees can be explained, or rather explained away, by either the no reference view, or its sibling view that there is no self-reference in central cases of first-person thought. The no reference view is associated with Anscombe and the less extreme view with Wittgenstein. Anscombe offers a purely non-referential account of ‘I’, while Wittgenstein gives a non-factualist reading of the central cases<sup>1</sup>. To accept either Anscombe’s or Wittgenstein’s explanations of these guarantees would be to undo the conceptual connection between self-consciousness and self-reference that is required for a defence of the Body Claim. In this chapter I argue that neither Anscombe’s nor Wittgenstein’s positions are acceptable, and that the two no reference views should be rejected<sup>2</sup>.

Before going on to describe the guarantees in question, it is necessary to forestall certain objections by saying something about Wittgenstein exegesis. It is well known that a central aspect of Wittgenstein’s project is to avoid giving theories. As a result, in many areas there is not really any such thing as ‘Wittgenstein’s position’. Non-factualism about self-ascriptions is no exception<sup>3</sup>. Whilst there are passages in which Wittgenstein *appears* to advocate non-factualism, there are others in which he clearly indicates that examples must be treated individually on a case by case basis. Thus it would be a mistake to think that any of the forms of non-factualism to be discussed in this chapter are, strictly speaking, Wittgenstein’s view. Having said this, it will be simpler if I continue to speak in this way. What is interesting for the present purposes is the view that Wittgenstein’s remarks point towards, even if it is not eventually endorsed by him. Such an interest seems legitimate since there are

---

<sup>1</sup> Non-factualism is often referred to as expressivism. My reason for declining to adopt this terminology will come out in what follows.

<sup>2</sup> Many of the questions pursued in this chapter are discussed in Brinck, 1997, Ch.2. Whilst I am sympathetic to many of Brinck’s conclusions, my treatment of the issues differs markedly from hers.

<sup>3</sup> An account of Wittgenstein’s attitudes towards non-factualism can be found in Hacker, 1990. For a note of scepticism concerning whether Wittgenstein endorsed the non-factualist view see Wright, 1998. For the opposing view see Jacobsen, 1996.

philosophers, currently endorsing a number of non-factualist positions, who are clearly influenced by Wittgenstein but who eschew his mistrust of theories.

The first-person guarantees in question are the following: (i) Guaranteed reference – ‘I’-thoughts cannot fail to refer, (ii) First-person authority – in central cases of ‘I’-thoughts we are guaranteed against making a mistake concerning the predicate which is self-ascribed, (iii) Immunity to error through misidentification – in central cases of ‘I’-thoughts we are guaranteed against making certain mistakes concerning who it is that we know to be such and such. Each of these guarantees requires a certain amount of explanation. I devote less time to guaranteed reference and first-person authority and more time to immunity to error through misidentification (IEM), for the reason that IEM is not generally that well understood. The following is partly an attempt to remedy this state of affairs.

## 2.2 Guaranteed Reference and First-Person Authority

Thoughts whose function it is to refer to objects can fail to do so, and can do so for different reasons. First, a thought can fail to refer because there is nothing for it to refer to, its intended object does not exist. This presumably is true of thoughts about Vulcan, or when one thinks about *that pink elephant* whilst hallucinating. Of these cases, some will say that the subject has thought a thought which fails to refer, others will say that the lack of referent means that no thought is really being thought<sup>4</sup>. Whatever one says of such cases, it would seem that the same is not possible with ‘I’-thoughts. It cannot be the case that one’s ‘I’-thought fails to refer due to the intended referent’s not existing.

A second way in which a thought can fail to refer is by there being too many objects. If I intermittently see two people, taking them to be one and the same person, Bert, I may go on to think, ‘Bert is here’. Further, if both Bert’s are here it is plausible to think that I have failed to refer to either<sup>5</sup>. Similarly, if I see the head of snake *a* and the tail of snake *b* poking out from either side of a rock, my thoughts about *that snake* will fail to refer to either. It would seem that the same is not possible with ‘I’-thoughts.

---

<sup>4</sup> For an object-dependent view of this kind see Evans, 1982.

<sup>5</sup> I ignore the possibility that this may count as a case of so-called ‘divided reference’. Divided reference might be thought to occur when the causal links governing a singular thought are deviant. If it is correct to characterise such cases, then it is plausible to think that ‘I’-thoughts will be immune to the phenomenon of divided reference. On divided reference see Kvat, 1989.

'I'-thoughts cannot fail to refer through an excess of objects<sup>6</sup>. On the face of it, the fact that 'I'-thoughts are guaranteed against these kinds of reference failure is something for which we should be able to provide an explanation. What this explanation might be is something to which we shall return.

Another phenomenon associated with first-person thought for which an explanation is certainly required is first-person authority<sup>7</sup>. First-person authority is what keeps in business those philosophers working on the problem of self-knowledge. First-person authority is a certain guarantee that subjects have concerning some mental self-ascriptions. 'I'-thoughts which self-ascribe beliefs, desires, wishes, emotions, sensations, intentions and so on, are known to be true by the subject with a certain authority that other kinds of thought do not possess. In standard cases of introspective self-knowledge – cases which do not involve irrationality or inference from observed behaviour – it cannot be that a subject sincerely thinks 'I believe that P' but has made a mistake and that in reality they believe that Q, or that not-P. Brute error is ruled out in such cases. Indeed, for a great many mental properties *P*, the person who is best placed to know whether or not I instantiate *P* is myself<sup>8</sup>. If I sincerely think, 'I am in pain', then the chances are that I am in pain. There may even be some mental properties of which I have absolutely infallible knowledge, but authority needn't amount to infallibility. Emotions provide a nice example. For the most part, the person best placed to say which emotion I am in is myself. But this is not always the case. Sometimes my emotions are hidden from me. I may, for example, discover that I am jealous by noting my reaction to some situation. This is useful for providing a negative characterisation of first-person authority, for in such a situation I lack the usual authority.

Adequately accounting for first-person authority has not proved to be an easy task, but some account is clearly required. Other people have epistemological access to our mental lives only through our observable behaviour. But this does not appear to be

---

<sup>6</sup> Lewis, 1976, offers an account of personal identity which makes room for the possibility that there can be two thinkers of one 'I'-thought (see especially, *Postscripts*). This is not the place to discuss his theory, but I assume that he would assent to a version of the self-reference view relativised to person-stages.

<sup>7</sup> It might be claimed that first-person authority is not relevant to the issue of first-person thought, since first-person authority (self-knowledge) need not involve the *I* concept. For instance, Hossack writes, "There are many creatures that can know by self-knowledge when they are in pain, even though they may not have a concept of the self" (Hossack, 2002, p.164). But what is it that these creatures are supposed to know? In describing the relevant knowledge we must say that they know when *they* are in pain, and this occurrence of 'they' is standing in for the *I* concept, see Casteñeda, 1966. It seems to me that if creatures are to be attributed self-knowledge they must be thought of as, at the very least, having something like the kind of non-conceptual self-consciousness described in Bermudez, 1998.

<sup>8</sup> As will become clear in the sections that follow, this characterisation of first-person authority in terms of knowledge begs the question against the non-factualist. At this point it is innocent enough, however.

the case with regard to the epistemological access that we have to our own minds. First-person authority seems to rest on something like a direct access to our own minds. Is this really the case? And if so, how is such access to be understood? Again, this explanation is something to which I shall return.

### 2.3 Immunity to Error Through Misidentification

It is often claimed that there are a range of self-ascriptions which are immune to error through misidentification relative to the first-person pronoun (IEM). What this means, and exactly which self-ascriptions are properly classed as IEM, is a topic hotly disputed. Some claim that only mental self-ascriptions are IEM, others claim that some physical self-ascriptions are IEM. The phenomenon of IEM was first brought to our attention by Wittgenstein who, in *The Blue Book* claims that,

There are two different cases in the use of the word 'I' (or "my") which I might call "the use as object" and "the use as subject". Examples of the first kind of use are these: "My arm is broken", "I have grown six inches", "I have a bump on my forehead", "The wind blows my hair about". Examples of the second kind are: "*I* see so and so", "*I* hear so and so", "*I* try to lift my arm", "*I* think it will rain", "*I* have toothache". One can point to the difference between these two categories by saying: The cases of the first category involve the recognition of a particular person, and there is in these cases the possibility of error, or as I should rather put it: The possibility of an error has been provided for...It is possible that, say in an accident, I should feel a pain in my arm, see a broken arm at my side, and think it is mine, when really it is my neighbours...On the other hand, there is no question of recognising a person when I say I have toothache. To ask "are you sure that it is *you* who have pains?" would be nonsensical. (Wittgenstein, 1958, pp.66-7)

Wittgenstein is here pointing out that a certain class of self-ascriptions are immune to a particular kind of error<sup>9</sup>. The same phenomenon is pointed to by Strawson in *The Bounds of Sense*. There he claims that

When a man (a subject of experience) ascribes a current or directly remembered state of consciousness to himself, no use whatever of any criteria of personal identity is required to justify his use of the pronoun 'I' to refer to the subject of that experience. (Strawson, 1966a, pp.164-5)<sup>10</sup>

---

<sup>9</sup> Wittgenstein suggests that it is a misunderstanding of this *grammatical* feature of 'I' which tempts many towards the Cartesian view that "the real I lives in my body" (1958, p.66). A similar suggestion is made in Strawson, 1966a, pp164-6.

<sup>10</sup> Also see Bennett, "I can wonder whose stomach is rumbling, but not whether it is my neighbour or myself who is embarrassed." (Bennett, 1966, p.109).

As elucidated by Wittgenstein and Strawson, the notion of IEM is left rather vague. It is Shoemaker who first gives us a precise definition. He writes,

to say that a statement '*a* is  $\phi$ ' is subject to error through misidentification relative to the term *a* means that the following is possible: the speaker knows some particular thing to be  $\phi$ , but makes the mistake of asserting '*a* is  $\phi$ ' because, and only because, he mistakenly thinks that the thing he knows to be  $\phi$  is what '*a*' refers to (Shoemaker, 1968, pp.7-8)

One thing that clearly must be added to this definition is relativity to grounds. It is obvious that a judgement may be IEM when based on one kind of ground, but subject to error when based on another. To see this, compare introspection with talking to one's analyst as ways of discovering one's mental states. Through introspection it is impossible for me to discover that *someone* resents their mother, and mistakenly think that it is *I* who resent my mother. Through talking to one's analyst, on the other hand, it is quite possible for me to discover that *someone* resents their mother, and mistakenly think that it is me. So, judgements are IEM relative to certain grounds.

So according to Shoemaker's definition, the self-ascription 'I have a headache' is IEM relative to introspection, since it is impossible for me to know, via introspection, that some particular person has a headache but mistakenly think that it is I who have it. Coming to know, via introspection, that someone has a headache is nothing less than coming to know that I have a headache<sup>11</sup>. Shoemaker's explanation of how it is that such judgements are immune to error is similar to that of both Wittgenstein and Strawson. It has, he believes, to do with the fact that such self-ascriptions are not based on an identification of the subject. It is impossible to misidentify oneself because there is no identification of oneself to go wrong.

This point is picked up by Evans in his treatment of IEM<sup>12</sup>. Immunity to error is a direct consequence of, what Evans calls, identification-freedom. According to Evans, to entertain a singular proposition about an object, the thinker must know which object the thought is about. One way this can be achieved is by means of an identification. For example, the thought '*a* is *F*' might be the result of the belief that '*b* is *F*' and the belief

---

<sup>11</sup> Campbell, 1999b, claims that the phenomenon of thought insertion shows that judgements based on introspection need not be IEM. This claim is sensibly disputed by Gallagher, 2000. A structurally similar debate concerning whether the phenomenon of anarchic hand syndrome shows that self-ascriptions of action based on proprioception are not IEM can be found in Marcel, 2003, and Peacocke, 2003.

<sup>12</sup> Evans, 1982, Chs.6 & 7.

that ‘ $a = b$ ’. Paraphrasing Evans, if knowledge of the truth of the singular proposition ‘ $a$  is  $F$ ’ can be seen as the result of knowledge of the truth of a pair of propositions ‘ $b$  is  $F$ ’ and ‘ $a = b$ ’ then we can call it ‘identification-dependent’. The latter proposition ‘ $a = b$ ’ is its identification component. If a singular proposition is identification-dependent, the thinker knows which object it is about in virtue of an identification. If, on the other hand, a thought does not rest on an identification component, then it is identification-free.

IEM, as defined by Shoemaker, follows from identification-freedom for the following reason: if a judgement is subject to error through misidentification, it must be possible for the thinker to make a mistake in the identification-component upon which it rests. But an identification-free judgement is not based on an identification-component, therefore it cannot be subject to error through misidentification, and so must be immune<sup>13</sup>. This close relationship between IEM and identification-freedom forms a common core running through the accounts of IEM given by Wittgenstein, Strawson, Shoemaker and Evans. We can articulate this common core as follows<sup>14</sup>:

**The proposition ‘ $a$  is  $F$ ’ is immune to error relative to grounds  $G$  iff every possible judgement that  $a$  is  $F$  (based on grounds  $G$ ) either:**

**(1) does not rest on two beliefs of the form ‘ $b$  is  $F$ ’ and ‘ $a = b$ ’.**

**or**

**(2) rests on beliefs of the form ‘ $b$  is  $F$ ’ and ‘ $a = b$ ’, but ‘ $a = b$ ’ is not rationally doubttable.**

This definition captures the way in which IEM is a consequence of identification-freedom<sup>15</sup>. According to this definition, if a judgement is identification-free, when

---

<sup>13</sup> Pryor, 1999, argues that IEM does not follow from identification-freedom. This is due to his idiosyncratic definition of IEM. I argue against Pryor’s definition of IEM in the Appendix.

<sup>14</sup> My definition of IEM treats it as a property of propositions. However, throughout this section I have often spoken of individual judgements as IEM. Whilst this is a useful way of speaking, it may strike one as odd, since it is unclear in what sense individual judgements are *immune* to error – they are either errors or they are not. The claim that a judgement  $J$  is IEM should be understood as shorthand for the claim that (i)  $J$  expresses a proposition which is IEM relative to  $G$ , and (ii)  $J$  is based on  $G$ .

<sup>15</sup> Evans’ definition of IEM is actually slightly different. He defines it as follows, “a judgement is identification-free [and thus immune to error] if it is based upon a way of knowing about objects such that it does not make sense for the subject to utter ‘Something is  $F$ , but is it  $a$  that is  $F$ ?’, when the first component expresses knowledge which the subject does not think he has, or may have, gained in any other way.” (1982, pp.189-190). This last constraint is not included in my definition of IEM. This should not worry us, since it is only necessary to introduce such a constraint if we rely on the idea of a judgement making sense for a subject. I avoid this locution, and intend Evans’ and my definitions to be equivalent.

based on grounds G, it will be IEM. The necessity for condition (2) can be seen if we consider a judgement which is identification-dependent but which the identification upon which it is based is not rationally doubtable. It is plausible to think that this will display immunity to error, since the indubitable character of the identification component will ensure that errors of misidentification will not occur. If this is right, identification-freedom and immunity to error could come apart.

Leaving such possible cases aside, we can think of IEM and identification-freedom as two sides of the same coin. We can also see how IEM relates to the functional role of 'I'-thoughts. In chapter one I said that 'I'-thoughts are liable to be affected by the various special ways that we have of gaining knowledge about oneself. These special ways clearly include introspection. Coming to believe, via introspection, that someone has a headache will dispose me to believe that I have a headache. By now this point is very familiar. The 'special ways' of gaining knowledge of oneself that play a part in the functional role of 'I'-thoughts just are those ways of gaining knowledge that lead to self-ascriptions that are IEM.

Simply pointing out that IEM is a consequence of identification-freedom is insufficient as an explanation of IEM. For we have yet to be given an explanation of identification-freedom. As we shall see, Wittgenstein suggests an explanation of IEM/identification-freedom in terms of a non-factualist account of self-ascriptions. Before addressing this suggestion, however, it is necessary to dispense with the objection that we have not characterized IEM correctly.

#### **2.4 Another Brand of Immunity to Error Through Misidentification?**

Some philosophers have defined immunity to error through misidentification in a way significantly different from the account given above. For instance, Wright says that,

A claim made on a certain kind of ground involves immunity to error through misidentification just when its defeat is *not* consistent with retention of grounds for existential generalization (Wright, 1998, p.19)<sup>16</sup>

This should be made clearer by example. Suppose I judge 'I was in Scotland last week', basing this upon my apparent memories of being in Scotland. However, I am then

---

<sup>16</sup> This different notion appears in discussions by Hamilton, 1995, Wright, 1998, Pryor, 1999, and Campbell, 1999a & 2002, Ch.5. In his, 1999b, Campbell appears to have something more like the standard definition of IEM in mind.

informed that my apparent memory of being in Scotland last week is, in fact, a quasi-memory. A quasi-memory being a state subjectively indistinguishable from a memory but which need not derive from the past history of the subject him or herself. If I have a quasi-memory of doing *x*, it may have been me who did *x*, or it may have been someone else. As such, memory is a sub-species of quasi-memory<sup>17</sup>. In this situation, I will no longer be justified in believing that *I* was in Scotland last week, but I will be justified in making the general judgement '*Someone* was in Scotland last week'. I can, as it were, retreat from the singular proposition to the general proposition, using the original grounds plus the piece of evidence that undercut my original judgement. The fact that I can do this, claims Wright, shows that my original judgement was not immune to error. On the other hand, consider the introspectively-based judgement 'I have a headache'. If someone (incredibly!) undercuts my grounds for thinking that *I* have a headache, it is arguably not possible for me to retreat to the judgement '*someone* has a headache'. This shows that my judgement 'I have a headache' is immune to error.

Now all of this is compatible with the view that propositions based on memory are IEM as I have defined it. It may well be that memory-based propositions rest on no identity component, yet their grounds can be retained as grounds for existential generalization. This is enough to show that the kind of immunity to error which Wright et al have in mind really is something new. We might come up with another name for it then: for want of something more eloquent, I shall call it IEM\*.

IEM\* is a rare phenomenon. Much more so than its proponents have thought. Suppose that, on the normal perceptual grounds, I form the judgement 'I see a canary'. I am then informed that really, my visual impression as of a canary is hallucinatory, and is caused by someone else's seeing a canary. Whenever they see a canary they press a button which causes me to suffer a canary-illusion. In this case I will be justified in retreating to the existential proposition 'someone sees a canary'<sup>18</sup>. Thus, there are at least some mental self-ascriptions which are not IEM\*<sup>19</sup>.

What, then, is IEM\* tracking? It is tempting to think that IEM\* is the result of a judgement's being infallibly known or its being groundless. A judgement is IEM\* when it is impossible for its grounds to be undercut in such a way as to leave open the

---

<sup>17</sup> The notion of a quasi-memory was first introduced by Shoemaker, 1970.

<sup>18</sup> This seems to me enough to cast very serious doubt on Prior's suggestion that Shoemaker is primarily concerned with IEM\*. See Pryor, 1999, pp.286-88; see Shoemaker, 1968, p.8, for his use of the canary example.

<sup>19</sup> Wright, 1998, Campbell, 1999a, and Pryor, 1999, all assume that this sort of present-tense mental self-ascription is IEM\*.

possibility of the corresponding existential judgement. But it seems that for any judgement whatsoever, if it is possible to undercut its grounds, it is possible to undercut them in such a way as to leave open just such a possibility. We can formulate this as a general thesis: any self ascription ‘I am *F*’ which it is possible to undercut, can be undercut by the information that one suffers the illusion that one is *F* in just those cases when someone else is *F*. This allows one to retreat to the existential proposition ‘*someone* is *F*’. So IEM\* arises in just those cases where a proposition’s grounds cannot be undercut at all. This will turn out to be the case if the proposition is infallibly known, or if it is based on no grounds whatsoever.

But if IEM\* just flags groundlessness or infallibility then it looks as if it defines too narrow a class of self-ascriptions to be of much independent interest. This is especially clear given that what many philosophers are looking to immunity to error to do is to capture that class of self-ascriptions seen as somehow central to self-consciousness. The class of first-person thoughts which is central to self-consciousness has traditionally included self-ascriptions of experiences. But the proposition ‘I can see a canary’, as we have seen, is not IEM\*. This might be conceded but it be pointed out that the judgement ‘I seem to see a canary’ is nevertheless IEM\*. And it is judgements of *this* kind that are really expressive of self-consciousness. This seems fair enough. But as long as we agree that IEM\* is tracking infallibility, there is scope for arguing that the introspectively-based self-ascriptions ‘I have a pain’ or ‘I am judging that P’ are not IEM\*. This will be the case if, for example, it is possible to mistake an itch for a pain, or to mistake an imagining for a judging.

To spell this out in the case of ‘I am in pain’: suppose you experience what you take to be a pain. Now suppose someone reliably tells you that you are not really experiencing a pain but an itch, and that this has been caused by someone else’s suffering a real pain. Whenever they experience a pain, they press a button which causes you to feel a pain-like itch. In this case, you will be justified in retreating from the singular proposition ‘I have a pain’, to the general proposition ‘someone has a pain’ whilst retaining the same grounds. If this describes a possible situation, and it certainly seems to, then ‘I am in pain’ is not IEM\*. But, of course, there is little reason to think that the self-ascription of sensations is not somehow central to self-consciousness<sup>20</sup>.

---

<sup>20</sup> If you are sceptical about the possibility of mistaking pains for itches etc. see the sorites-style argument in Williamson, 2000, Ch.4.

My argument that IEM\* is not a property of judgments such as ‘I am in pain’ may strike some as somewhat dubious. Of course, in these cases, the subject will be justified in retreating to the general proposition, but isn’t this justification all packed into the piece of defeating evidence? This feature might seem objectionable. Can’t we simply reformulate our definition of IEM\* so that the justification that the subject has for retreating to the general proposition must already be there in the justification for the original singular judgement? This would rule out my examples as counterexamples to the claim that fallible mental self-ascriptions can be IEM\*. The problem with this suggestion is that it becomes impossible to distinguish IEM\* from IEM. If, in the justification that I have for the singular judgement ‘I am *F*’ I am able to discern two elements, one of which can be discarded and the other retained, leaving me justified in thinking ‘Someone is *F*’ then it looks very much as though that original judgement was based on the information that *a* was *F* and that  $I = a$ , i.e. it was based on an identification. And now the difference between the two forms of IEM is entirely eroded. Thus, if the proponent of IEM\* wishes that notion to be distinguished from the standard conception of IEM, he or she had better not make this change to the formulation.

I take it, then, that IEM\* is not a particularly interesting phenomenon in its own right. Of course, the fact (if it is one) that some judgements are infallible or groundless is philosophically interesting, but we knew that already. Infallibility is just a special case of first-person authority. As we have already allowed that the phenomenon of first-person authority requires an explanation, it seems unnecessary to speak of IEM\*, and it is particularly misleading to treat it as if it were a distinct phenomenon<sup>21</sup>.

## 2.5 The No Reference View

Now that we have a decent view of the phenomena of guaranteed reference, first-person authority and IEM, we can look at the claim that they are best explained by versions of the no reference view. I begin with guaranteed reference. Since guaranteed reference holds for all ‘I’-thoughts its explanation must be similarly general. It will have to be explained by the extreme version of the no reference view according to which ‘I’-thoughts never refer. The no reference view, rather than explaining guaranteed

---

<sup>21</sup> This is somewhat unfair to Wright, 1998, who suggests that IEM\* is not a particularly interesting phenomenon, since it is a trivial consequence of groundlessness. But whilst Wright correctly points out that groundlessness is sufficient for IEM\*, the point of the present discussion is to argue that groundlessness (or infallibility) is both necessary and sufficient for IEM\*.

reference, explains it away. According to the no reference view, self-ascriptions do not refer to the person who thinks them, they do not refer at all. It is no surprise then that they cannot fail to refer, referring is not their function. The no reference view explains the guarantee that 'I'-thoughts have against reference failure, not by explaining how they are guaranteed to refer, but by showing that there is no question of success or failure here.

The no reference view is defended and put to this very use by Anscombe<sup>22</sup>. Anscombe presents a host of considerations against the view that 'I' is ever a referring expression. Although Anscombe's arguments are stated at the level of language, I take it that she would suppose them to have analogues at the level of thought. The central case against treating 'I' as a referring expression is the following: first, Anscombe takes it that if 'I' is a referring expression, then it is an expression whose reference is, in a certain sense, guaranteed. Second, she thinks that if 'I' is a referring expression it must be understood on the model of either a proper name, a demonstrative, or an abbreviation of a definite description. Further, each of these kinds of referring expression requires, what Anscombe calls, a 'conception' by means of which it reaches its referent. This conception must be such as to explain the guaranteed reference of 'I'. However, the argument continues, no satisfactory conception can be specified for 'I': either it fails to deliver up guaranteed reference, or it succeeds but only by delivering an immaterial soul. Since there are no immaterial souls, it follows that 'I' cannot be understood on the model of proper name, demonstrative or definite description. It follows from this that 'I' is not a referring expression. The analogous conclusion at the level of thought is that 'I'-thoughts are not self-referential, and this is the extreme version of the no reference view.

There are a number of ways in which Anscombe's argument can be challenged. It might be argued (as in Evans, 1982, Ch.7) that 'I' follows the model of 'here' or 'now'. Or Anscombe's claim that every referring expression must be associated with a conception might be challenged. Finally, it could be pointed out that guaranteed reference could be explained by some means other than the conception associated with 'I'. I will not directly challenge Anscombe's argument. But I will give reason to think

---

<sup>22</sup> Anscombe, 1975. Discussions of Anscombe's arguments can be divided into those who, broadly speaking, agree with her position, and those who do not. Those who agree include, Clarke, 1978, Kenny, 1979, and Malcolm, 1979. Those who disagree include, White, 1979, Hamilton, 1991, Brandon, 1994, pp.552-561, McDowell, 1998, and Harcourt, 2000. The Wittgensteinian/Anscombian view is criticised in Glock and Hacker, 1996. However, they suggest that 'I' is "best seen as a degenerate or limiting case of a referring expression" (1996, pp.95).

that her conclusion must be false. If an argument can be mounted successfully against the more moderate ‘no reference in central cases’ view, then Anscombe’s position will fall with it. For if it can be shown that central cases of ‘I’-thoughts are self-referential, it has already been shown that the extreme view is false. The rest of this chapter constitutes my argument against the more moderate version of the no reference view, and so also against Anscombe’s position. It will also be argued that there is another, more satisfactory, explanation of guaranteed reference, thus rendering the no reference view explanatorily redundant.

## 2.6 Non-Factualism

Whilst guaranteed reference is a feature of all ‘I’-thoughts, first-person authority and IEM are features of the central cases only. An adequate explanation of these phenomena must reflect this. One view, inspired by certain remarks of Wittgenstein, is that such an explanation is to be found in non-factualism about self-ascriptions. The idea is that non-factualism will allow us to explain away both of these phenomena. The further claim is that the truth of non-factualism about mental self-ascriptions entails that central cases of first-person thought are not self-referential.

Non-factualism is best seen as the rejection of certain assumptions of what we might call the standard picture<sup>23</sup>. According to the standard picture, mental self-ascriptions are utterances that express propositions, are straightforwardly true or false, function as assertions, and the acceptance of which is constituted by a belief in the proposition expressed. Indeed, they are not only believed but known to be true in some, yet to be explained, direct manner. To clarify, let us take as an example of a mental self-ascription, an utterance of, ‘I have a headache’. The standard picture claims that this self-ascription expresses the proposition that I have a headache, and that its expressing this proposition is (at least partly) constitutive of its having the meaning it does. The self-ascription represents me as having a headache. If I have a headache the self-ascription will be true, if I have no headache it will be false. To say that this self-ascription is an assertion is to say that it is put forward as a truth, in uttering it I commit myself to its truth. When I accept a self-ascription I believe the proposition that is

---

<sup>23</sup> In thinking about non-factualism I have been greatly helped by discussions with Mark Kalderon, who also kindly allowed me to read a draft of his forthcoming book, *Moral Fictionalism*. Much of what I have to say in the rest of this chapter is directly influenced by that work.

expressed. When I utter, 'I have a headache', I do so because I accept that sentence, that is I believe that I have a headache.

Non-factualism rejects all of these assumptions of the standard picture. The non-factualist about mental self-ascriptions claims that they are non-representational, they do not express propositions, rather their meaning is fixed by the fact that they express the first-order state that they *appear* to represent. As a consequence, they do not function as assertions, and acceptance of a mental self-ascription is not belief. Some quotations from Wittgenstein will help us to get a feel for this view. Directly after he first introduces the notion of IEM (quoted at length above), Wittgenstein offers a non-factualist explanation,

And now this way of stating our idea suggests itself: that it is impossible that in making the statement "I have a toothache" I should have mistaken another person for myself, as it is to moan with pain by mistake, having mistaken someone else for me. To say "I have pain" is no more a statement *about* a particular person than moaning is. (Wittgenstein, 1958, p.67)

Wittgenstein makes essentially the same point in *Philosophical Investigations*. It is worth quoting Wittgenstein at length here, as it is useful to see the connections between this and the material already quoted,

"When I say 'I am in pain', I do not point to a person who is in pain, since in a certain sense I have no idea *who* is"...I did not say that such-and-such a person was in pain, but "I am..." Now in saying this I do not name any person. Just as I don't name anyone when I *groan* with pain...What does it mean to know *who* is in pain? It means, for example, to know which man in this room is in pain: for instance, that it is the one sitting over there with the fair hair, and so on.—What am I getting at? At the fact that there is a great variety of criteria for personal '*identity*'...Now which of them determines my saying that '*I*' am in pain? None. (Wittgenstein, 1953, §404)

Wittgenstein's suggestion is that utterances in which 'I' is used as subject (are IEM) are expressive rather than representational. Furthermore, since he seems to take mental self-ascriptions to be IEM quite generally, it would seem that he is offering a non-factualist account of mental self-ascriptions. Another passage which might be taken to support such a view is the following,

Here is one possibility: words are connected with the primitive, the natural, expressions of the sensation and used in their place. A child has hurt himself and he cries; and then adults talk to him and teach him exclamations and, later, sentences. They teach him new pain behaviour..."So you are saying that the

word ‘pain’ really means crying?”—On the contrary: the verbal expression of pain replaces crying and does not describe it. (Wittgenstein, 1953, §244)

Here we see Wittgenstein suggesting that mental self-ascriptions are expressive rather than representational<sup>24</sup>. If this view is correct, the weak version of the no reference view is vindicated. For, on this account, the central cases of self-consciousness do not represent the subject as being a certain way. But, if this is the case, the central cases of self-consciousness are not self-referential.

It is extremely natural for a worry to arise at this point. The conceptual connection between self-consciousness and self-reference upon which the Body Claim rests is supposed to hold at the level of thought. But non-factualism is standardly taken to be a thesis concerning language. After all, its central claims concern assertion and expression. Why should the proponent of the Body Claim be worried about the idea that certain linguistic self-ascriptions are not self-referential? The simple answer is that it would be extremely odd to think that an utterance of ‘I have a headache’ does not express the proposition that I have a headache, whilst the corresponding thought *does* express that proposition. Non-factualism about mental self-ascriptions seems to sit more comfortably with the view that self-ascriptions in thought are non-representational, that they do not represent the thinker as being a certain way. If non-factualism is true the linguistic expression of central cases of self-consciousness are not referential, and this pushes us towards the view that the first-person thoughts themselves are not self-referential. As such, the defender of the Body Claim does have a reason to be concerned about non-factualism.

The central claim of non-factualism is the non-representational character of mental self-ascriptions.

**The Non-Representation Thesis:** Mental self-ascriptions do not express propositions which represent the subject as having some mental property.

---

<sup>24</sup> Recall my qualms about interpreting Wittgenstein. That Wittgenstein is probably not making a general claim is supported in the following passage, “A cry is not a description. But there are transitions. And the words “I am afraid” may approximate more, or less, to being a cry. They may come quite close to this and also be *far* removed from it.” (Wittgenstein, 1953, Part II, p.189). Also, note the emphasis in this following, “‘I believe...’ throws light on my state. Conclusions about my conduct can be drawn from this expression. So there is a *similarity* to expressions of emotion, or mood, etc.” (Wittgenstein, 1953, Part II, p.191). Furthermore, Wittgenstein’s denial that the word ‘pain’ really means crying, quoted in the main text, suggests that he does not accept non-factualism as I have been formulating it.

This is the point that Wittgenstein makes when he denies that the sentence ‘I have a pain’ is a statement about a particular person, that it describes the subject. Given certain plausible assumptions, several other theses follow from the non-representation thesis. One such assumption is that truth and falsity is primarily a property of propositions. Sentences are true or false in virtue of expressing true or false propositions. It follows that non-factualism is committed to the following claim,

**The Non-Truth-Evaluable Thesis:** Mental self-ascriptions are neither true nor false in the standard sense<sup>25</sup>.

Another plausible assumption is that both belief and assertion are relations between a speaker and a proposition<sup>26</sup>. When a subject asserts that P he or she holds the assertion relation to the proposition that P. Similarly, when a subject believes that P he or she holds the belief relation to the proposition that P. It follows that the non-factualist is committed to the following two theses,

**The Non-Assertion Thesis:** The assertive utterance of a mental self-ascription is not the assertion of a proposition.

**The Non-Cognitivist Thesis:** The acceptance of a mental self-ascription is not belief in a proposition.

The non-factualist’s acceptance of the non-cognitivist thesis should presumably be thought to involve a commitment to the view that mental self-ascriptions are not known. These are all negative claims that the non-factualist makes concerning mental self-ascriptions, but the non-factualist is committed to some positive claims also. In particular, the non-factualist claims that rather than representing the subject as being a certain way, mental self-ascriptions are expressive of the first-order state they appear to represent. For example, an utterance of ‘I have a headache’ does not represent the utterer as having a headache, rather it is an expression of the headache that is had. But here we have reached a confusion in terminology. I have been speaking of sentences expressing propositions, this is a semantic relation. But now we are told that utterances

---

<sup>25</sup> This is consistent with self-ascriptions being true\* or false\* in some non-standard sense.

<sup>26</sup> For the purposes of this chapter I disregard the objections raised by Lewis, 1979, to the view that belief is a relation to a proposition.

express mental states, and this is a pragmatic relation. Indeed, most philosophers who have discussed this type of non-factualism have referred to it as *expressivism*. Kalderon (forthcoming) suggests the term ‘convey’ for this pragmatic relation, but he defers to established usage. To avoid confusion, I will use the somewhat cumbersome ‘c-express’ and its cognates, to refer to the relation that non-factualists claim holds between mental self-ascriptions and the first-order mental states they appear to represent. The non-factualist thus holds,

**The C-Expression Thesis:** Mental self-ascriptions c-express (convey) the first-order mental states they appear to represent.

There is one final claim made by the non-factualist which is central to the position. This is a claim about the meaning of mental self-ascriptions. Standardly, when a sentence expresses a proposition the meaning of that sentence consists (at least partly) in its expressing that proposition. Of course, since the non-factualist denies that mental self-ascriptions express propositions, this view is unavailable. The non-factualist claims, rather, that the meaning of mental self-ascriptions consists in their c-expressing the first-order states they appear to represent. Thus,

**The Semantic Thesis:** The meaning of a mental self-ascription does not (even partly) consist in its expressing a proposition. Rather, the meaning of a mental self-ascription consists entirely in its c-expressing the first-order state it appears to represent.

These then are the central commitments of non-factualism about mental self-ascriptions. We can now ask how it is that non-factualism explains the phenomena of first-person authority and IEM.

Above, I characterized first-person authority by saying that if I sincerely judge ‘I am in pain’, then the chances are that I am in pain. The non-factualist focuses on this way of saying what first-person authority is. The non-factualist’s acceptance of the non-truth-evaluable thesis and the non-cognitive thesis prevents him or her from saying either of the following things: (i) Sincere mental self-ascriptions are, to a greater or lesser extent, guaranteed to be true, (ii) Each person has ‘direct’ and authoritative knowledge of their own mental states. As such, there is a sense in which non-factualists

deny self-knowledge—the position is a dissolution rather than a solution to problems arising from the phenomenon of first-person authority. Hence, Wittgenstein’s remark,

I can know what someone else is thinking, not what I am thinking.

It is correct to say “I know what you are thinking”, and wrong to say “I know what I am thinking.”

(A whole cloud of philosophy condensed into a drop of grammar.) (Wittgenstein, 1953, Pt.II, p.222)

Non-factualists claim that mental self-ascriptions are not, in the normal circumstances, knowledgeable. Thus, there is no problem in explaining how we come about such knowledge. Since self-ascriptions are c-expressive rather than representational, there is nothing there to know. All there is to be explained is that a sincere utterance of, say, ‘I am in pain’ is a highly reliable indicator that the utterer is in pain. This, it seems, is well explained by the c-expression thesis. For, self-ascriptions are, when sincere, c-expressions of the first-order states they appear to represent. But if a sincere utterance is a c-expression of a state, then it is a highly reliable indicator of the existence of that state in the same way that, for example, a groan is a highly reliable indicator of the existence of a pain state. It seems, then, that non-factualism does provide a good deflationary account of first-person authority.

We must also ask how successful non-factualism is in explaining the phenomenon of IEM. To put matters roughly, IEM is a feature of self-ascriptions that employ no criteria of identity of the subject. They are self-ascriptions that do not rest on an identification-component. And now it seems that non-factualism, if true, explains this phenomenon. For mere c-expressions of mental states employ no criteria of identity of the subject (or anything else), for the reason that they do not represent the subject (or anything else). Thus, if self-ascriptions are given a non-factualist reading, they will turn out to be IEM. As Wittgenstein puts it,

it is impossible that in making the statement “I have a toothache” I should have mistaken another person for myself, as it is to moan with pain by mistake, having mistaken someone else for me. (Wittgenstein, 1958, p.67)

There is one worry, however, concerning the power of non-factualism to explain IEM. Non-factualism is a thesis about mental self-ascriptions only. Nobody would claim that physical self-ascriptions are c-expressive of the states that they appear to describe. It follows that non-factualism can only give a *full* explanation of IEM if IEM is a feature

of mental self-ascriptions only. However, some have claimed that IEM is a feature of some physical self-ascriptions also<sup>27</sup>.

There are at least three ways in which the non-factualist could respond to this worry. First, they could simply deny that any physical self-ascriptions are IEM, on the grounds that they are not c-expressive. Second, they could attempt to show how physical self-ascriptions ‘borrow’ their IEM from mental self-ascriptions, upon which they are based<sup>28</sup>. Third, they could agree that non-factualism only gives a partial explanation of IEM but, it must be said, an explanation of the most central cases of IEM. Whichever way the non-factualist goes, it is clear that he or she has at least the beginnings of an explanation of IEM. When this is coupled with the explanation of first-person authority, non-factualism does have some explanatory success.

Before moving on to the reasons for thinking non-factualism an unacceptable view, I will say something about the c-expression thesis. I have already given one reason for thinking that *expressivism* is not a good label for the non-factualist view (the ambiguity of the word ‘express’), but there is another reason. This second reason is that the standard picture of mental self-ascriptions as representational and so on, can happily accept the c-expression thesis. In fact, there is a sense in which the proponent of the standard picture is more of an expressivist than is the non-factualist.

To use a phrase taken from Wittgenstein, a c-expression of a mental state ‘throws light on my state’. An interesting, and less metaphorical, analysis of c-expression is given by Taylor, 1979. Taylor regards a c-expression as a bodily manifestation of a state which reveals that state by communicating it to others<sup>29</sup>. This view correctly entails that crying out c-expresses pain, smiling c-expresses joy, etc.

Non-factualism claims that mental self-ascriptions c-express the first-order states they appear to represent. Is this plausible? Let  $\phi$  range over the psychological verbs of which non-factualism is supposed to hold<sup>30</sup>. What, if anything, does an utterance of ‘I  $\phi$  (that) P’ c-express? Let us first take belief: ‘I believe that P’. It seems reasonable to say that this utterance is a bodily manifestation of the belief that P which reveals that state by communicating it to others. This sits well with the fact that uttering

---

<sup>27</sup> The first to claim this was Evans, 1982, Ch.7.

<sup>28</sup> Wright, 1998, §2, suggests something like this.

<sup>29</sup> This is a very rough gloss on Taylor’s much more nuanced view. I would not wish to commit myself to Taylor’s view of c-expression. I do have concerns, especially with the ‘communication’ aspect. This is especially clear for my own purposes, since I am primarily interested in thought, not language.

<sup>30</sup> The following is easily modified to allow for those self-ascriptions where the mental term appears as an adjective, e.g. ‘I am in pain’ or ‘I am happy’.

'I believe that P' c-expresses a commitment to the truth of P. But what is a commitment to the truth of P if not a belief that P? If this is correct, the self-ascription of belief will, as the non-factualist maintains, indeed c-express the mental state it appears to represent.

However, the standard picture can accept this point. There is no obvious reason to think that accepting the c-expression thesis undermines the claim that mental self-ascriptions are representational, etc. Furthermore, unless one is already committed (as is the non-factualist) to the view that acceptance of 'I believe that P' is not belief, there appears little reason for denying that 'I believe that P' is a bodily manifestation of the second-order belief that I believe that P which reveals *that* state by communicating it to others. The proponent of the standard picture can consistently maintain that acceptance of 'I believe that P' is belief and that an utterance of that sentence is c-expressive of that (second-order) belief. Thus, the proponent of the standard picture can claim that an utterance of 'I believe that P' c-expresses *both* the state of believing that P, and the state of believing that one believes that P. It is in this sense that the proponent of the standard picture is being more of an expressivist than the non-factualist, for the non-factualist denies that an utterance of 'I believe that P' is c-expressive of the belief that one believes that P as, according to the non-factualist, there is no such belief.

What about self-ascriptions of other kinds of mental predicates, such as 'I hope that P', 'I want such-and-such' and so on. In each case, it is highly intuitive to say that the self-ascription c-expresses the first-order state it appears to describe. Indeed, it is reasonable to think that the c-expression thesis is quite generally true for mental self-ascriptions. Utterances of mental self-ascriptions are c-expressive of the first-order states they appear to describe. We can see that not only is this compatible with the standard picture of mental self-ascriptions, but the proponent of the standard picture is also perfectly free to claim that utterances of mental self-ascriptions are also c-expressive of the second order belief in/knowledge of the proposition that I  $\phi$  (that) P. Thus, self-ascriptions are doubly c-expressive.

It might be objected that it is implausible to think that a single piece of (verbal) behaviour could be c-expressive of two distinct states<sup>31</sup>. On this view, since self-ascriptions typically c-express the first-order states they appear to represent, they cannot c-express any second-order state. But this view has little to recommend it. My violent outburst may well express both my rage and my insecurity. To defend such a view, we would need to be given an account of what c-expression is which precluded one

---

<sup>31</sup> Jacobsen, 1996, makes such a claim.

utterance from c-expressing two distinct states. But it seems likely that any such account would be a defective account of c-expression.

## 2.7 The Frege-Geach Problem

There are a family of arguments, made familiar by Geach, 1965, against the form of non-factualism I have been discussing. This has become known as the Frege-Geach problem. Roughly speaking, the Frege-Geach problem consists in pointing out that the set of sentences that the non-factualist claims are c-expressions (in this case, mental self-ascriptions) can occur in contexts in which they cannot be expressions of anything. The first difficulty for the non-factualist involves inferences. Consider the following instance of modus ponens:

- (1) If I believe that P, then I should believe that Q.
- (2) I believe that P.

Therefore,

- (3) I should believe that Q.

On the face of it this is a valid argument. But in order for an instance of modus ponens to be valid both the antecedent of the first premise and the second premise must have the same meaning. If they differ in meaning, the argument will be invalid due to equivocation. According to non-factualism, the second premise does not express a proposition, rather its meaning consists in its c-expressing the belief that P. The problem arises when we notice that the antecedent of the first premise does not c-express the belief that P. The first premise can be sincerely uttered by someone who fails to believe that P, indeed by someone who believes not-P. If the antecedent of the first premise does not function as a c-expression of the belief that P, then its meaning cannot consist in its c-expressing the belief that P. So according to the non-factualist this argument is invalid. But this is disastrous as it casts doubt upon the coherence of a large portion of our capacity for practical and theoretical reasoning, for it is plausible to think that the kind of reasoning displayed in the above example occurs in much of our rational decision making<sup>32</sup>.

---

<sup>32</sup> The importance of this kind of first-personal reasoning is stressed in Burge, 1998.

The second difficulty for the non-factualist about mental self-ascriptions involves negation. According to the non-factualist mental self-ascriptions have no truth-values or, at least, are neither true nor false in the standard sense. But, in this case, what are we to make of an utterance of 'It is not the case that I believe that P'? Negation is an operation which reverses truth-values. If the mental self-ascription has no truth-value, then its negation is meaningless. But of course this is not the case. Therefore the occurrence of 'I believe that P' which appears after the negation operator must differ in meaning from the standard assertive utterance. To this it might be responded that 'It is not the case that I believe that P' c-expresses the belief that not-P. But this is false, for in uttering 'It is not the case that I believe that P' I might be giving voice to my agnosticism regarding P. A better response is to argue that there is a kind of truth, truth\*, according to which c-expressions can be true\*. But the defender of this view owes us an account of truth\* and of negation\*.

A third difficulty for non-factualism is that it appears to make intersubjective conversation about one's mental states near impossible. If my utterance of 'I believe that P' does not express the proposition that I believe that P, then it means something different from your utterance of 'You believe that P'. We are talking past each other. Worse, when I utter 'I believe that P' it is impossible for you to believe what I said, since I have not expressed a proposition and belief is a relation between a subject and a proposition. It might be argued that my utterance c-expresses my belief that P, and so in believing me, you bear the belief relation to the proposition that P. But this is wrong, as it should be possible for you to assent to what I said without thereby believing P. A theist can agree with my utterance of 'I believe that God does not exist'. The problem is that, according to non-factualism, first and third-person utterances of mental predicates differ in meaning. As Heal puts it, on this view, "our psychological talk fragments into two unrelated halves" (Heal, 2002, p.3).

This semantic ambiguity concerning the first and third-person also infects the present and past-tense. Presumably the past-tense 'I believed that P' expresses the proposition that I believed that P. It cannot be a c-expression of my belief that P, for I may no longer have that belief. But if I can assert the proposition that I believed that P, why can I not assert the same thing at a time which requires the present-tense formulation? Wittgenstein considers and rejects this line of thought in the following,

“But surely ‘I believed’ must tell of just the same thing in the past as ‘I believe’ in the present!”—Surely  $\sqrt{-1}$  must mean just the same in relation to  $-1$ , as  $\sqrt{1}$  means in relation to  $1$ ! This means nothing at all. (Wittgenstein, 1953, Part II, §X)

But the analogy is a weak one. The difference is that the truth of ‘I believe that P’ at  $t_1$  entails the truth of ‘I believed that P’ at  $t_2$ . No such relation holds between  $\sqrt{-1}$  and  $\sqrt{1}$ . One might be tempted to say that ‘I believed that P’ c-expresses my memory of believing that P. But this still leaves us with semantic ambiguity: the meaning of the present-tense self-ascription is constituted by its c-expressing a belief, the meaning of the past-tense self-ascription is constituted by its c-expressing a memory.

Finally, an objection has recently been put forward to non-factualism based on minimalism about truth<sup>33</sup>. Minimalism about truth is the view that all there is to truth is the disquotation schema: ‘P’ is true if and only if P. According to minimalism, P can be replaced by any grammatically suitable sentence, and mental self-ascriptions are clearly grammatically suitable sentences. Thus, self-ascriptions have truth conditions. For example, ‘I am in pain’ is true if and only if I am in pain. It follows that, at least, the non-truth-evaluable thesis is false. Without this, it is suggested, non-factualism collapses. For, the falsity of the other components of non-factualism follows from the falsity of the non-truth-evaluable thesis<sup>34</sup>.

These objections to non-factualism about mental self-ascriptions are formidable. And if non-factualism about mental self-ascriptions is false, it cannot explain first-person authority or IEM. If non-factualism is false, we need to come up with alternative explanations. It is worth, then, seeing if the view can be salvaged. One way of doing this would be to attempt to develop non-factualism in such a way as to accommodate the claim that self-ascriptions have the semantic structure they appear to have. This response usually attempts to draw a parallel between self-ascriptions and performatives. For reasons that will become clear, I shall call this view *Fictionalism*<sup>35</sup>.

---

<sup>33</sup> See, in particular, Wright, 1992.

<sup>34</sup> Of course, this final argument relies on the truth of minimalism. One might want to reject it, then, by rejecting the claim that the disquotation schema tells us all there is to know about truth. As Jacobsen, 1996, points out, this option is not open to Wittgenstein, who appears to accept minimalism.

<sup>35</sup> It might be thought that non-factualism could be salvaged by attempting to construct a ‘logic of attitudes’, thereby providing a quasi-realistic theory of self-ascriptions. This is the option, suggested in the main text, of attempting to give sense to notions such as truth\* and negation\*. On quasi-realism, see Blackburn, 1984, Ch.6. I do not have the space to explore this option here, but I have some reservations. First, the complications of constructing an alternative logic are to be avoided if possible. If a simpler theory is in the offing it should be tried and tested first. Second, there are serious doubts about the success of the quasi-realist programme, see Hale, 1993.

## 2.8 Fictionalism

Whilst the Frege-Geach problem spells the end of non-factualism, some of its central theses remain unscathed. For example, although he is no non-factualist Wright (1998, p.36) denies that the Frege-Geach problem has any real force against a suitably sophisticated view<sup>36</sup>. He regards the problem as showing that self-ascriptions such as ‘I believe that P’ have truth-evaluable content, but fall short of showing that they assert anything. He cites performatives as a model: ‘I promise to pay’ has truth-evaluable content, yet in uttering the sentence I do not assert that content. On this picture of performatives, ‘I promise to pay’ is true if and only if I promise to pay, and is therefore true every time sincerely uttered. This explains the fact that ‘I promise to pay’ can be meaningfully negated, appear in instances of modus ponens etc. However, in uttering ‘I promise to pay’, I have not asserted that I promise to pay. Rather, I have promised to pay.

Looking back to the six theses of non-factualism, this position rejects the non-representation thesis, the non-truth-evaluable thesis and the semantic thesis. That is, mental self-ascriptions *do* express propositions, and so can be true or false, and their meanings are (at least partly) constituted by their expressing those propositions. Rejecting these theses allows us to accept all the morals of the Frege-Geach point. However, we are still in a position to accept the non-assertion thesis, the non-cognitive thesis and the c-expression thesis. That is, the assertive utterance of a mental self-ascription is not assertion, the acceptance of a mental self-ascription is not belief, and mental self-ascriptions c-express the first-order mental states they (actually) represent. This view I will refer to as fictionalism.

The fictionalist view takes over that part of the standard picture of mental self-ascriptions which pertains to semantics. The meanings of mental self-ascriptions are just as we pre-theoretically expect them to be. However, fictionalism takes over that aspect of non-factualism which pertains to pragmatics. Although mental self-ascriptions are representational, accepting one does not amount to believing it, and uttering one does not amount to asserting it. The same can be said, it is suggested, about performatives<sup>37</sup>.

---

<sup>36</sup> Variations of this view are defended in Jacobsen, 1996, Bar-On, 2000, and Bar-On & Long, 2001. The issues are subtly discussed in Falvey, 2000, who endorses the c-expression thesis but not the other aspects of non-factualism or fictionalism.

<sup>37</sup> This is, of course, in opposition to Austin, “None of the [performative] utterances cited is either true or false: I assert this as obvious and do not argue it” (1962, p.6). The view that performatives are, in fact, truth-evaluable is expressed in Lewis, 1970, p.224, and Davidson, 1968, p.107.

It is not clear to me, however, that the model of performatives is really helpful here. Is it true that when I utter, ‘I promise to pay you on Tuesday’ I neither assert nor believe the proposition expressed? Here is an intuitive argument against such a view. If, on hearing my utterance, Bill says ‘Smith said that he promised to pay on Tuesday’, Bill has said something true. But surely this means nothing other than ‘Smith asserted that he promised to pay’. To assert something is to put it forward as true, and when I say ‘I promise to pay on Tuesday’ I put it forward as true that I promise to pay on Tuesday. Furthermore, there seems no obvious reason to think that in accepting the sentence ‘I promise to pay on Tuesday’ I do not believe that proposition expressed. What could justify such a claim?

This may not convince many. But there is another reason to think that, whilst there is an analogy between mental self-ascriptions and performatives, the analogy is far from complete. For, for the most part, performatives are self-constituting. By this I mean that by sincerely uttering a performative sentence one *thereby* makes it true. It is my saying that I promise to pay on Tuesday that makes it true that I promise to pay on Tuesday. This self-constituting feature of performative utterances is not shared by mental self-ascriptions. Certainly sincere utterances of mental self-ascriptions are invariably true, but this is not because they make themselves true. Uttering ‘I believe that P’ does not make it true that I believe that P. So much appears to be clear<sup>38</sup>.

A better model is fictional discourse (hence the label). When I say, ‘Holmes is a great detective’, I have uttered a sentence which is representational, expresses a proposition which (at least in part) constitutes its meaning, and is either true or false. However, I have not asserted that proposition, and in accepting the sentence I do not believe the proposition it expresses. In fictional discourse we do not assert or believe the propositions expressed by the sentences we utter for the reason that we do not think them literally true. We speak ‘within the fiction’. Here, then, is a better model for understanding fictionalism about mental self-ascriptions.

But the fact, if it is one, that fictionalism is internally consistent and not worried by the Frege-Geach problem is not the end of the matter. For we have yet to see whether it manages to adequately explain first-person authority and IEM. Fictionalism offers less satisfactory an explanation of first-person authority than does non-factualism. Non-factualism rejected the idea that self-ascriptions were representational, and that they were truth-evaluable in the standard sense. Thus, says the non-factualist, there is

---

<sup>38</sup> Pace Heal, 2002.

nothing to be known (hence the non-cognitivist thesis). First-person authority is not a matter of direct knowledge of mental states. Rather, it is simply that self-ascriptions are reliable indicators of a person's mental state. Fictionalism, on the other hand, allows that mental self-ascriptions are both representational and truth-evaluable. But now self-ascriptions express propositions which might, or might not, be known. Instead of likening mental self-ascriptions to groans of pain, of which it makes no sense to ask whether they are known, the fictionalist admits that they have content; that they say something. The fictionalist is now in a position to say that when mental self-ascriptions are sincere, they will, for the most part, be true. But, this is a less satisfactory explanation than the one given by non-factualism, for there is a further question to be answered, 'Are mental self-ascriptions known?'. If the answer is yes, we have rejected the fictionalist non-cognitivist thesis and must explain how it is that they are known. If the answer is no, we can legitimately ask why not. Both non-factualism and fictionalism deny self-knowledge, but non-factualism has good reason—it makes no sense to say that mere c-expressions are known—fictionalism provides no such reason.

How good an explanation can fictionalism give of IEM? Non-factualism explained IEM by pointing out that mere c-expressions of mental states employ no criteria of identity of the subject. Since self-ascriptions do not refer to the person who utters them, there is no need for any criteria of identity to be employed in securing that reference. But the fictionalist claims that self-reference does occur in self-ascriptions. We have now lost our explanation of IEM. The fact that self-ascriptions are c-expressive does not help to show how it is that they employ no criteria of identity, for they are no longer mere c-expressions. The fact that the acceptance of a mental self-ascription is not belief is no help here. It remains that self-ascriptions are representational. This is true for both language and thought. In thinking the thought 'I have a headache', I have thought a thought with self-referential, representational content. The fictionalist claim is that I do not bear the belief relation to this thought. But if there is reference we are owed an explanation of how it is that errors of misidentification do not occur. Fictionalism does not provide such an explanation.

This last point should spell the end of our concern with fictionalism as a putative obstacle to the Body Claim's reliance on a conceptual connection between self-consciousness and self-reference. Fictionalism is not a version of the no reference view. According to fictionalism, mental self-ascriptions are representational, they represent the subject as being a certain way. But to do this, they must refer to the subject. Thus,

even if fictionalism about mental self-ascriptions did provide a good explanation of first-person authority and IEM, it gives us no reason to doubt the conceptual connection between self-consciousness and self-reference.

## 2.9 Explaining the First-Person Guarantees

If, as we have seen reason to believe, the view that there is no self-reference in the central cases of first-person thought is false, then the view that 'I'-thoughts are *all* non-referential is false. This means that none of the first-person guarantees with which we began have been adequately explained. In the present section I suggest how we might proceed in giving such explanations.

Let me begin with guaranteed reference. Guaranteed reference is explained by the self-reference rule. What needed explanation was that 'I' cannot fail to refer through there being either too many or too few objects. The self-reference rule easily explains this. The rule, which determines the reference for each and every utterance of 'I', states that an utterance of 'I' refers to the person who utters it. When I utter a sentence containing 'I', my intended reference is *myself*<sup>39</sup>. But since 'I' cannot be uttered by anyone who does not exist, 'I' cannot fail to refer due to the lack of existence of its intended object.

Similarly, the self-reference rule explains why 'I' is immune to reference-failure through there being too many objects. A single token of 'I' cannot be uttered by two distinct people. If two people utter 'I', even if they are in the same place at the same time, there will be two utterances of 'I', each one referring to the person who produced it. So the self-reference rule determines that 'I' cannot fail to refer through there being too many objects. Everything that goes for language in this context also goes for thought. These explanations are, I claim, far more natural and satisfactory than either non-factualism or Anscombe's no-reference thesis. Of course this explanation relies on the self-reference rule and, as was pointed out in chapter one, there exist some reasons for thinking that there may be exceptions to the self-reference rule. But, of course, any reason for thinking that there are exceptions to the self-reference rule is also a reason for thinking that there are exceptions to guaranteed reference. This does not impugn the explanation.

---

<sup>39</sup> Anscombe, 1975, thinks this problematic because circular. For criticism see, O'Brien, 1994. This issue does not matter for the present purposes.

The task of explaining first-person authority is huge and this is not the place to take it on. I will, therefore, limit myself to explaining why non-factualism and fictionalism point us in the wrong direction. In their explanations of first-person authority, both non-factualism and fictionalism rely on the c-expression thesis. Earlier I said that self-ascriptions are, when sincere, c-expressions of the first-order states they appear to represent. Now if a sincere utterance is a c-expression of a state S, then it is a highly reliable indicator of the existence of S. Just as a groan is a highly reliable indicator of the existence of a pain state. But there are at least two reasons to think that this is, after all, not an adequate deflationary account of first-person authority. First, as Heal has argued, “for all we know there will be a fair number of false positives” (Heal, 2002, p.9). The c-expression thesis does not rule out the possibility that a subject may frequently sincerely utter ‘I believe that P’, when no belief is being expressed. Just as it is conceivable that a subject may, in good faith, be disposed to groan with pain even though they feel none, so a subject may, in good faith, be disposed to utter ‘I believe that P’ even though they do not. The second worry concerns the concentration on language rather than thought<sup>40</sup>. Let us suppose that the c-expression thesis can provide a deflationary explanation of the first-person authority we enjoy over our utterances of mental self-ascriptions. We still have not explained first-person authority in its entirety, for we are strongly inclined to say that a subject knows what he or she is thinking, desiring, feeling, etc. even if he or she displays no overt behaviour. The c-expression thesis is a thesis about (linguistic) behaviour. But first-person authority is a phenomenon which seems not to be adequately captured by an account of how subjects behave. First-person authority is something which is available to a subject who is not in a position to express any of his or her states behaviourally. These criticisms apply to both the non-factualist and fictionalist accounts of first-person authority. Whatever the correct account of first-person is, it should not rely on the c-expression thesis.

The explanation of IEM is to be found in the postulation of, what I shall call, ‘single-object faculties’. Discussions of IEM tend to concern the question whether such-and-such faculty for gaining knowledge gives rise to judgements which are IEM. Thus, there have been debates as to whether judgements based on memory or bodily-awareness are IEM. What needs to be made clear is how one might go about claiming that self-ascriptions based on bodily-awareness or self-locating perceptions are IEM. That is, we need an answer to the question, what is it that one needs to claim about a

---

<sup>40</sup> See Wright, 1998, p.37.

given faculty for gaining knowledge about oneself, *f*, in order to secure the claim that self-ascriptions based on *f* are IEM?

One way of doing this is to claim that *f* is necessarily a way of gaining knowledge of only one object<sup>41</sup>; what we might call a ‘single object faculty’. Take the case of bodily-awareness. If one holds that bodily-awareness is necessarily awareness of one’s own body<sup>42</sup>, it seems that one will have all the materials required for a defence of the claim that self-ascriptions based on bodily-awareness are IEM. If bodily-awareness is necessarily an awareness of one’s own body, it follows that one cannot be aware, via bodily-awareness, of the body of another. So, it will be impossible to know that someone’s legs are crossed without thereby knowing that one’s own legs are crossed. Looking back at the definition of IEM, we can see that, on this conception of bodily-awareness, the proposition that ‘*b* is *F*’ could not be distinct from the proposition ‘*a* is *F*’, for there will be only one object with which one is acquainted through bodily-awareness. If (based on bodily-awareness) I am in a position to make the singular judgement ‘*a* is *F*’ then, ex hypothesi, either it is based on no identification, or the identification upon which it is based is indubitable. This for the reason that it is not possible to be aware of and to hold singular beliefs about a body, perceived via bodily-awareness, unless that body is one’s own.

Exactly similar remarks can be made for other faculties. For instance, the claim that introspection-based mental self-ascriptions are IEM is secured by the, surely correct, claim that introspection is necessarily a way of coming to know the contents of one’s own mind. The claim that memory is necessarily a way of gaining knowledge of one’s own past secures the claim that memory-based self-ascriptions are IEM. And so on. Arguing that a faculty is a single object faculty is a way of arguing that judgements based on that faculty are IEM. Nothing more seems to be required. Thus, whether or not one thinks that judgements based on a faculty, *f*, are IEM depends entirely on what one takes to be the correct theory of *f*.

This ties in with the way in which Wittgenstein and Strawson originally characterized IEM, as not requiring the application of any criteria of personal identity. For, if a faculty is a single object faculty, then coming to know via that faculty that someone is *F* just is coming to know that I am *F*. There is no need for me to apply any criteria of personal identity for me to discover who the person is that I know to be *F*.

---

<sup>41</sup> See Campbell, 1999a, pp.92-3.

<sup>42</sup> As is argued (disregarding some complications) in Martin, 1995.

I have been arguing against non-factualism about self-ascriptions. In doing so, I have defended the intuitively compelling view that, at least, the central cases of first-person thought are self-referential. The phenomena of guaranteed reference, first-person authority and immunity to error through misidentification do not tell against self-reference. The conceptual connection between self-consciousness and self-reference is intact.

## 3. Russell's Principle

### 3.1 The Ambiguous Status of Russell's Principle

The defender of the Body Claim is committed to there being a tight conceptual connection between self-consciousness and self-reference. We have seen off one type of challenge to this connection. There is, however, another obstacle to be overcome. This comes in the form of those accounts of singular thought which adopt, what has become known as, Russell's Principle. I shall be arguing against Russell's Principle and thereby defending the close connection between self-consciousness and self-reference.

The role played by Russell's Principle in the dialectic of the Body Claim is, however, ambiguous in the extreme. On the one hand, acceptance of the principle can readily lead one to think that central cases self-consciousness need not involve self-reference. In this respect Russell's Principle represents an obstacle to be overcome by the defender of the Body Claim. On the other hand, those philosophers who have defended the Body Claim, and related views, have often used Russell's Principle as a premise in their arguments *for* the Body Claim! In this respect the rejection of Russell's Principle deprives proponents of the Body Claim of one of their greatest weapons. In general there has been a failure to appreciate this crucial ambiguity. Once we have a clear view of the situation, we begin to understand the rather unusual way in which Evans fits in to the picture. For the position occupied by Evans in *The Varieties of Reference* is very naturally thought of as, at the very least, closely allied to the Body Claim. Yet, as we shall see in this chapter, Evans' view in fact poses a serious threat to the defender of the Body Claim, for Evans's view of singular thought entails that self-consciousness need not involve self-reference. This is a direct result of the absolutely central importance accorded by Evans to Russell's Principle. In chapters four and five we shall see how Russell's Principle has been used to support later stages of arguments for the Body Claim. The remainder of this chapter will be to defend my claim that Russell's Principle Provides a serious challenge to the defender of the Body Claim. This will be carried out through a detailed analysis of Evans' account of singular thought. At the end of the chapter, I go on to argue against Russell's Principle.

### 3.2 What Russell's Principle Is

Russell articulated his principle of acquaintance in several places. Here is one,

*Every proposition which we can understand must be composed wholly of constituents with which we are acquainted...*The chief reason for holding this principle true is that it seems scarcely possible to believe that we can make a judgement or entertain a supposition without knowing what it is that we are judging or supposing about...It seems to me that the truth of this principle is evident as soon as the principle is understood (Russell, 1910-11, pp.159-160).

Acquaintance is an epistemic notion; being acquainted with an object involves having a certain kind of knowledge of it. This knowledge is 'knowing which object it is that one is thinking about'. In line with his epistemological views, Russell himself thought that the only items with which we are acquainted in the requisite way were sense data, universals, our own mental states and, possibly, ourselves. Understood as involving these epistemological limitations, the principle of acquaintance has little to recommend it<sup>1</sup>. It follows from Russell's official position that either the proposition that London is the capital of England can be analysed into a proposition making reference only to sense data etc. or that it cannot be understood.

But it seems reasonable to think that the heart of the principle of acquaintance can be divorced from Russell's own epistemological strictures. This is the strategy adopted by Evans. Evans claims that to be in a position to entertain a singular thought about *a* one must know which thing *a* is, but he thinks that one can have this kind of knowledge about, amongst other things, everyday physical objects. The kind of knowledge that Evans believes is required for one to know which object one is thinking about is, what he calls, 'discriminating knowledge'. Evans calls his version of the principle of acquaintance 'Russell's Principle' and I shall follow him in this. He writes,

a subject cannot make a judgement about something unless he knows which object his judgement is about...the knowledge which it [Russell's Principle] requires is what might be called *discriminating knowledge*: the subject must have the capacity to distinguish the object of his judgement from all other things. (Evans, 1982, p.89)

---

<sup>1</sup> See, for example, Sainsbury, 1979, pp.26-41, who argues against Russell's position.

This means that unless I have such a capacity (to distinguish the object of my thought from all other things), I cannot think about it. I shall sometimes refer to the act of distinguishing an object from all other things as ‘identifying’ it. Now there are, according to Evans, three ways in which I can satisfy Russell’s Principle with respect to any given object: (i) I can identify it by perceiving it, (ii) I can identify it by having the ability to recognise it were I to perceive it, and (iii) I can identify it through my knowing distinguishing facts about it.

As we shall see, Evans’ account of how it is that we identify ourselves is a version of (i). That is Evans thinks that the way in which we distinguish ourselves from all other things, and thereby satisfy Russell’s Principle with respect to ourselves (and are thereby able to entertain singular thoughts about ourselves) is by perceiving ourselves. To understand how an adherence to Russell’s Principle leads Evans to the view that a subject might be self-conscious and yet incapable of self-reference, we need to look at some of the other aspects of his account of singular thought. Once we see how this general account deals with first-person thought, we shall see that his position leaves open the possibility of, what I shall call, massive reference failure. We will also see that the possibility of massive reference failure is something that the proponent of the Body Claim must deny. My view is that this should be done via a rejection of Russell’s Principle itself.

### **3.3 Evans on Singular Thought**

We need now to take a relatively lengthy detour through some other aspects of Evans’ account of singular thought. In particular we must introduce the Generality Constraint, the distinction between fundamental Ideas, and the notion of an information-based thought. The reasons for doing this will become clear in the following sections.

Alongside Russell’s Principle, Evans further claims that singular thoughts must conform to the Generality Constraint<sup>2</sup>. Evans takes thoughts to be structured: they are, he thinks, composed of concepts, conceived of as abilities to think of objects. So, for instance, the thought ‘Bruce is bald’ is composed of the concept *Bruce* and the

---

<sup>2</sup> Travis, 1994, argues forcefully against the Generality Constraint. Unfortunately, I do not have the space to discuss his arguments here.

concept of *being bald*. Anyone who can think that thought, possesses those two concepts. One of these is a concept of an object, one is a concept of a property. Following Evans<sup>3</sup>, we can call the first of these an Idea of an object. An Idea of an object is a concept of an object, that is, an ability to think of that object. Anyone able to think 'Bruce is bald', has both an Idea of *Bruce* and the concept of *being bald*.

We can now state the Generality Constraint. If a subject has an Idea of an object, *a*, then that subject has the conceptual resources to entertain the thought '*a* is *F*' for every concept *F* which he or she possesses (and is categorially applicable). Similarly, if a subject has the concept *F*, that subject has the conceptual resources to entertain the thought '*a* is *F*' for every object of which he or she has an Idea.

The Generality Constraint requires us to see the thought that *a* is *F* as lying at the intersection of two series of thoughts: the thoughts that *a* is *F*, that *a* is *G*, that *a* is *H*, ..., on the one hand, and the thoughts that *a* is *F*, that *b* is *F*, that *c* is *F*, ..., on the other. (Evans, 1982, p.209)

The Generality Constraint is, of course, an elaboration of a point made by Strawson in *Individuals*<sup>4</sup>. There, Strawson claims that one does not have the idea of a predicate unless one has the idea of something which can be affirmed of a *range* of individuals.

Next we must move on to a distinction, made by Evans, between fundamental and non-fundamental Ideas. To understand this distinction between fundamental and non-fundamental Ideas of objects, we need another piece of Evans' terminology: 'the fundamental ground of difference'. Every object has a fundamental ground of difference. The fundamental ground of difference of an object is that which differentiates that object from other objects of the same kind. So, for instance, the fundamental ground of difference (at a time) for a physical object, is its spatial location plus its sort (statue, lump of clay etc.). That is, physical objects are differentiated from each other by their spatio-temporal location plus sort. Stating the fundamental ground of difference of a given kind is just (or is at least closely related to) stating the criterion of identity for objects of that kind.

A fundamental Idea of an object is an Idea of that object as being the possessor of the fundamental ground of difference that it actually has. Evans writes that,

---

<sup>3</sup> Who claims to be following Geach, 1957. But see Geach, 1986.

<sup>4</sup> Strawson, 1959, p.99.

one has a *fundamental Idea* of an object if one thinks of it as the possessor of the fundamental ground of difference which it in fact possesses. (Evans, 1982, p.107)

An example of a fundamental Idea of an object would be one based on the perception of the object in one's immediate environment. If I see a cat in front of me and think 'that cat is black', I will be thinking demonstratively of the cat as the possessor of the fundamental ground of difference it, in fact, has; namely as located in such and such a time and place, and as being of such and such a kind.

On the other hand, thinking of my cat, Tiddles, on the other side of the city, I will not have a fundamental Idea of him<sup>5</sup>. If I think 'Tiddles is black', I am not thinking of him as being located in the place in which he is, in fact, located. I don't know where he is at the present time. My Idea of Tiddles satisfies Russell's Principle because, were I to see him, I would recognise him.

Notice that the example given of a fundamental Idea exploits the demonstrative mode of identification. Indeed, demonstrative Ideas of physical objects are quite generally fundamental. On the other hand, the example given of a non-fundamental Idea exploits the recognition-based mode of identification. Both the recognition-based and the descriptive modes of identification give rise to non-fundamental Ideas.

What exactly are non-fundamental Ideas of objects? Obviously, in my thought about Tiddles I am not thinking of him as the possessor of the fundamental ground of difference which he in fact possesses. And this is true of non-fundamental Ideas in general. However, according to Evans, non-fundamental Ideas do exploit the fundamental level of thought.

Let  $\delta$  range over fundamental Ideas of objects. Evans thinks that I know what it is for the proposition 'Tiddles is black' to be true, because I know what it would be for the propositions ' $\delta$  is black' and ' $\delta = \text{Tiddles}$ ' to be true. So it turns out that my understanding of propositions containing non-fundamental Ideas of objects rests on my understanding of propositions containing fundamental Ideas.

The first of the above propositions, ' $\delta$  is black' is just a standard proposition involving a fundamental Idea of an object. The second, however, is new. The second proposition is a *fundamental identification*. A fundamental identification is an

---

<sup>5</sup> Tiddles is, to my knowledge, no relation of Geach's well known moggy 'Tibbles'.

identification between a fundamental Idea of an object and a non-fundamental Idea. It is the capacity to know what it would be for such fundamental identifications to be true that is distinctive of the ability to entertain thoughts containing non-fundamental Ideas of objects. Evans writes,

we can take the subject's Idea-of-the-object,  $a$ , to consist in his knowledge of what it is for an arbitrary proposition of the form  $\lceil \delta = a \rceil$  to be true (Evans, 1982, p.110)

But what does this mean? One thing it surely can't mean is that in order to have a non-fundamental Idea of an object, one must know its actual fundamental ground of difference. If it did mean this, there would be nothing to distinguish fundamental from non-fundamental Ideas. It seems to me that the most reasonable way of understanding non-fundamental Ideas is that they require that one knows the *sort* of fundamental ground of difference relevant to the kind of object one is thinking about. So, for instance, if one has a non-fundamental Idea of a physical object, one must know what it would be for it to be located somewhere. That is, one must have a ability to conceive of the very object of one's thought as located in such-and-such a place. Evans brings this out later, in his discussion of demonstrative identification,

In the case of a spatio-temporal particular, an adequate Idea of an object [i.e. either fundamental or non-fundamental] involves either a conception of it as the occupant of such-and-such a position (at such-and-such a time), or a knowledge of what it is for an object so identified to be the relevant object (Evans, 1982, p.149)

The final aspect of Evans' account of singular thought to which we must turn, is his notion of an information-based thought. Information-based thoughts are thoughts which employ Ideas that are governed by information gained in certain ways. Typically, these 'information-channels' are perceptual. For example, suppose I can see a cat (Evans, 1982, p.121). My thoughts about *that cat* will be governed by the content of my perception of it. If I see that the cat is black, I will be disposed to think that the cat is black, and so on.

Information-based thoughts are dependent upon the existence of such information-channels. That is, they require an information-link; a link between the object of one's thought and one's Idea of that object. If the information (or

misinformation) which partly comprises my Idea of an object actually derives from either more or less than exactly one object, then my attempts to think the relevant thoughts will fail. I will come back to this kind of ‘ill-groundedness’ later. Information-based thoughts include thoughts which employ both Ideas that are demonstrative, and Ideas that are recognitional, but not those that are purely descriptive. When a subject thinks about an object via a definite description, that subject’s thoughts need not be governed by any information about that object. As we shall see later, it is a moot point whether ‘I’-thoughts are information-based.

### 3.4 ‘I’-Ideas

In the present section I will outline Evans’ account of first-person thought. It is important to ask how first-person thought fits in to Evans’ general view of singular thought just described. It transpires that there is a difficulty with Evans’ view, centring around the issue of the distinction between fundamental and non-fundamental Ideas. This difficulty, which specifically concerns the possibility of non-fundamental ‘I’-Ideas, will be of importance when, in the next section, we come to discuss the possibility of massive reference failure.

‘I’-Ideas, in Evans’ terminology, are the self-conscious Ideas we have of ourselves, ‘I’-thoughts are thoughts which contain ‘I’-Ideas. Evans treats ‘I’-thoughts on a demonstrative model. They are, he claims, similar to ‘this’-Ideas and, in particular, ‘here’-Ideas<sup>6</sup>. This similarity shows up in the functional role of ‘I’-Ideas. This point should be familiar from chapters one and two.

First, ‘I’-Ideas feed into our dispositions to act in a special way<sup>7</sup>. If I believe that *I* am about to be attacked by a tiger, I will be disposed to run away. However, if I believe that Smith is about to be attacked by a tiger, I will only be disposed to run away if I also believe that I am Smith. This feature is common to both ‘I’-thoughts and ‘here’-thoughts. If I believe there is a tiger *here*, I will be disposed to run away.

---

<sup>6</sup> That ‘I’-Ideas are to be modelled on ‘here’-Ideas would have been Evans’ answer to Geach’s question: “*what* is it that on his view the pronoun ‘I’ is? What feature of ‘I’ has Anscombe failed to notice?” (Geach, 1986, p.535). What Anscombe, 1975, failed to notice was that ‘I’ can be treated on the model of ‘here’. Geach obscures the point, however, by giving the impression that Anscombe explicitly denies that ‘I’ follows the model of ‘here’. But there is no such denial in Anscombe’s paper, where the only demonstratives considered are ‘this’ and ‘that’.

<sup>7</sup> See, in particular, Perry, 1979.

On the other hand, if I believe that there is a tiger in my living room, I will only be disposed to run away if I also believe that my living room is *here*.

Second, 'I'-thoughts like 'here'-thoughts are disposed to be controlled by information gained in certain ways. Take 'here'-thoughts. If I see a tiger, I will be disposed to think 'there is a tiger here'. My 'here'-thoughts will be determined by the information I receive via the perception of my immediate environment. There is no need for me to identify the place at which the tiger is as *here*. Another way to put this is to say that 'here'-thoughts, when based on perception, are identification-free. That is, they don't rest on an identification of the presented location as *here*. This identification-freedom gives rise to the IEM of perception-based 'here'-thoughts.

So, 'here'-thoughts are disposed to be controlled by information gained via the perception of the immediate environment. Evans' claim about 'I'-thoughts is that they behave in a similar way, except that the information channels by which they are disposed to be controlled, are different. One such information channel is introspection. If I have a headache, I will be disposed to think 'I have a headache' or 'I am in pain'. Furthermore, this does not require an identification of the person who has the headache as me. The thought 'I have a headache' is identification-free, and this is what explains the fact that it is IEM.

According to Evans, much the same can be said of certain of the ways we have of gaining knowledge of our bodily properties. Specifically, our method of gaining knowledge of the state of our own bodies (bodily-awareness), and our method of locating ourselves relative to our perceived environment. So, the judgements 'I have crossed legs' or 'I am in front of a tree' are, when based on the appropriate grounds, both identification-free. It is in these facts that we can see the similarity between 'I'-thoughts and 'here'-thoughts which Evans uses to justify his treatment of 'I'-Ideas on the model of demonstratives. This similarity is put clearly in the following,

the Idea which one has of oneself involves the same kinds of elements as we discerned in the case of, say, 'here': an element involving sensitivity of thoughts to certain information, and an element involving the way in which thoughts are manifested in action. (Evans, 1982, p.207)

If what was claimed in chapter two is correct, these claims made by Evans can be secured by treating introspection, bodily-awareness and 'self-locating perception' as single object faculties.

I shall now go back to the general features of the theory of singular thought. For 'I'-thoughts are singular thoughts, and as such we should be able to apply Evans' general theory to them. We can begin by asking how it is that 'I'-thoughts satisfy Russell's Principle. Russell's Principle states that in order to make a judgement about an object, I must know which object my judgement is about. Three ways of satisfying this constraint are identified by Evans: I can currently perceive the object, I can recognise it, or I can know distinguishing facts about it. Evans' likening of 'I'-thoughts to demonstratives suggests that he has a version of the first in mind. That is, knowing which object my 'I'-thought is about is a matter of being able to perceive it, roughly speaking.

But this hasn't really told us very much yet. We need to know more about 'I'-Ideas. Specifically we need to know how it is that we identify ourselves. This necessity becomes acute when we admit that 'I'-Ideas conform to the Generality Constraint. We can grasp all sorts of propositions about ourselves, not just those which are controlled by the special information channels mentioned earlier. Since this is the case, there is more to be said about our 'I'-Ideas than just the bare facts that they are controlled by certain ways we have of gaining information about ourselves, and that they feed into dispositions to act in a special way. I can grasp the thought 'I will die'. In order to do this, according to Russell's Principle, I must know which object I am thinking about. And this means that I must be able to distinguish that object from all other things. But this is not done via those special information channels, mentioned earlier. So how is it that I am able to do this?

Perhaps a better way to approach this question is to ask whether the Idea I have of myself in grasping the thought 'I will die' is fundamental or non-fundamental. And, more generally, whether 'I'-Ideas are fundamental or non-fundamental. This is a vital question and, unfortunately, Evans is not explicit. Recall, one has a fundamental Idea of an object if one thinks of it as the possessor of the fundamental ground of difference it in fact has. So before we can answer the question of whether 'I'-Ideas are fundamental or non-fundamental, we need to know what the fundamental ground of difference for persons is.

What is the fundamental ground of difference for persons? This is a *metaphysical* question, and the answer isn't obvious. However, Evans isn't interested in going too far into the metaphysical question and he simply assumes an answer. The answer he proposes is as follows,

It seems to me clear that as we conceive of persons, they are distinguished from one another by fundamental grounds of difference of the same kind as those which distinguish other physical things, and that a fundamental identification of a person involves a consideration of him as the person occupying such-and-such a spatio-temporal location. (Evans, 1982, p.211)

This is a roundabout way of saying that persons are physical objects<sup>8</sup>. Not such a controversial assumption but, to digress a little, it does serve to dispel one false impression that certain of Evans' remarks are liable to give rise to. For instance, it has been suggested by Brewer, 1995, that Ch. 7 of *The Varieties of Reference* constitutes, in part, an argument against Cartesian dualism. And this interpretation is not without plausibility. After all, Evans does speak of certain of his views as being "the most powerful antidote to a Cartesian conception of the self" (Evans, 1982, p.220). It is true to say that Evans' intentions are to give a firmly anti-Cartesian account of self-reference, self-consciousness and self-identification. It would be a mistake, though, to suppose that this is to be done by *arguing* against Cartesian metaphysics. There simply is no sustained argument against the view that the fundamental ground of difference for persons should be spelt out with reference to souls. No doubt Evans *would* argue against this. But he doesn't. And this isn't really so surprising, since Ch. 7 isn't a work in metaphysics. Evans' isn't an anti-Cartesian argument, but an anti-Cartesian position<sup>9</sup>.

Returning to the point at hand, persons are individuated by their spatio-temporal location and the fact that they are persons. So, having a fundamental Idea of oneself would involve thinking of oneself as the person at such and such a spatio-temporal location. Are our 'I'-Ideas like this? Well yes, it would seem that in at least some cases they are. In particular, the special information channels mentioned earlier that give me knowledge of my bodily properties, namely bodily-awareness and my capacity to locate myself, are important here. For these are precisely the most

---

<sup>8</sup> Well, almost. I shall be suggesting in chapter six that a Cartesian dualist might want to claim that souls are individuated by the spatio-temporal location of the bodies to which they are connected.

<sup>9</sup> One might try to squeeze an anti-Cartesian argument out of Evans in the following way: on Evans' view, if Cartesian dualism were true, our 'I'-Ideas would be ill-grounded through there being too many objects, body and soul, (see the discussion of ill-groundedness in the following section), and so our 'I'-thoughts would fail to net any object. Thus, the subject would lack the ability to think 'I'-thoughts. This might be taken to be an argument against Cartesian dualism (or maybe an, indirect, argument against Evans' account of singular thought!). Of course, the dualist's reply will be that, contrary to appearances, we have two distinct types of 'I'-Idea: 'I<sub>soul</sub>'-Ideas, and 'I<sub>body</sub>'-Ideas.

immediate ways I have of coming to discover my location. When I judge, based on visual perception, that I am in front of a tree, I am thinking of myself as the person in such and such a location. So it seems open for Evans to say that, in thinking such a thought, I have a fundamental Idea of myself.

But this won't work for all 'I'-thoughts. For we can't forget about the Generality Constraint. In saying that 'I'-thoughts conform to the Generality Constraint, we are saying that I must be able to entertain thoughts such as 'I will die', and since these thoughts are not immediately controlled by the information channels that give me information of my own location, it is not obvious that these thoughts involve a fundamental Idea of oneself. At least this is what is suggested by Evans' treatment of this very example,

It is not wholly inaccurate to say that I grasp such an eventuality by thinking of myself in the way that I think of others; this is just another way of saying that the fundamental level of thought about persons is involved. But it is of course essential that I am aware that the person of whom I am so thinking *is myself*; certainly I must have in mind what it is for  $\lceil \delta \text{ is dead} \rceil$  to be true, for arbitrary  $\delta$ , but I must also have in mind what it is for  $\lceil \delta = I \rceil$  to be true. My thought about myself does satisfy the Generality Constraint; and this is because I can make sense of identifying a person, conceived from the standpoint of an objective view of the world, as myself. (Evans, 1982, p.210).

This looks very much as though Evans is saying that in thinking thoughts such as 'I will die' we employ non-fundamental Ideas of ourselves, for a fundamental identification (of the form  $\lceil \delta = I \rceil$ ) is only a requirement when non-fundamental Ideas are concerned. Possessing a non-fundamental Idea of an object does not require one to be in possession of the fundamental ground of difference of that object. So, if some 'I'-Ideas are non-fundamental, they do not require one to be in possession of one's own fundamental ground of difference at that time. All one needs to know is what it would be for one to be identical with a person, thought of using a fundamental Idea, i.e. one must know what it would be for  $\lceil \delta = I \rceil$  to be true<sup>10</sup>.

---

<sup>10</sup> It might be objected that 'I'-Ideas must be either fundamental or non-fundamental, but we cannot pick and choose as we please. It is a misunderstanding of Evans' account to think that some 'I'-thoughts could employ fundamental 'I'-Ideas, and some not. But what is an Idea but an ability to think of an object? Since this is the case, we should not be surprised that an Idea will differ according to the different circumstances in which it is employed. Currently I can see Tiddles my cat and so have an ability to think about him. My Idea of him is fundamental. But, suppose he runs away and I no longer know where he is. Of course I can still think about him, but my Idea will no longer be fundamental. I am suggesting that the same is true of 'I'-Ideas.

But now we arrive at a problem. How do non-fundamental 'I'-Ideas conform to Russell's Principle? 'I'-Ideas, based as they are on the model of demonstrative reference, satisfy Russell's Principle by being grounded in a perception of the object. In thinking of myself, I distinguish myself from all other things by my ability to locate myself in space. But if I have a non-fundamental Idea of myself, I need not be in possession of my actual fundamental ground of difference. But it is now difficult to see how I am able to distinguish myself from all other things. The only types of non-fundamental Ideas discussed by Evans are recognition-based Ideas and description-based Ideas. But Evans explicitly disavows a description-based account of 'I' (Evans, 1982, p.206), and surely a recognition-based model would be hopeless.

I take it that the answer to this problem is supposed to lie in the fact that, for many of one's 'I'-thoughts, one *can* distinguish oneself from all other things. In actual fact, one can so distinguish oneself, and this somehow makes it possible for one to entertain thoughts about oneself involving non-fundamental 'I'-Ideas. Evans tells us that,

to know what it is for  $\ulcorner \delta_i = I \urcorner$  to be true, for arbitrary  $\delta_i$ , is to know what is involved in locating oneself in a spatio-temporal map of the world. (Evans, 1982, p.211)

But knowing what is *involved* in locating oneself does not entail that one *can* locate oneself, and so does not entail that one can identify oneself, i.e. distinguish oneself from all other things. Thus, it seems that if this is all one must be able to do to have a non-fundamental Idea of oneself, then such thoughts must conform to Russell's Principle by being dependent, in some respect, on thoughts which involve a fundamental Idea of oneself. But how this might be cashed out remains unclear<sup>11</sup>.

Another way of raising the same point, a way which will lead us naturally on to the next section, is to consider the possibility of one's having the various special ways of gaining knowledge of oneself blocked. That is, the possibility of total sensory deprivation. It is worth asking what happens to a subject's capacity to entertain 'I'-thoughts when he or she is subjected to the sort of sensory deprivation discussed by Anscombe, 1975. This is a problem that O'Brien, 1995, has raised and she takes it

---

<sup>11</sup> Essentially the same difficulty is raised by Peacocke, 1983, pp.169-174, concerning perceptual demonstratives. For a discussion of both perceptual demonstratives and self-identification, see the exchange between McDowell, 1990, and Peacocke, 1991.

that Evans' position can't deal with the coherence of first-person thought in the sensory deprivation tank.

So here is the challenge to Evans: According to the theory, 'I'-Ideas must conform to Russell's Principle, i.e. we must know which object we are thinking about. In the case of 'I'-thoughts, this is achieved via our capacity to locate ourselves in a spatio-temporal map of the world. But we can imagine a subject in a sensory deprivation tank, with all relevant information channels blocked. It seems relatively plausible to think that this subject can think 'I won't let this happen again' (to use Anscombe's example). But how, it can be asked, does this thought conform to Russell's Principle? This is precisely a case in which the subject *can't* locate himself<sup>12</sup>. So we might think that Evans is committed to saying that the subject in the sensory deprivation tank is incapable of first-person thought. Evans denies this,

A subject may be amnesiac and anaesthetized, and his senses may be prevented from functioning; yet he may still be able to think about himself, wondering, for example, why he is *not* receiving information in the usual ways. But it would be...wrong to conclude from this that self-consciousness can be explained without reference to the various ways that subjects have of gaining knowledge about themselves...It is essential, if a subject is to be thinking about himself self-consciously, that he be *disposed* to have such thinking controlled by information which may become available to him in each of the relevant ways. (Evans, 1982, pp.215-6)

So, the anaesthetized subject can entertain 'I'-thoughts just in case *if he were* to regain the use of the normal information channels, then his 'I'-thoughts *would* be controlled by them. This move parallels one made in the case of 'here'. For, admits Evans, a subject can think of a place as 'here' whilst currently receiving no information from that place. In the case of 'here' this is because places are "so much thicker on the ground than objects, a subject cannot fail to have a single place as the target of his 'here'-dispositions" (Evans, 1982, p.169). As a result of this, Evans claims that

We have now what might be regarded as a limiting case of information-based thoughts: the subject's thinking is not necessarily controlled by a conception, but it rests upon an information-link in that the subject is so disposed that his thinking involving the Idea in question will be controlled by information yielded by the link *if any emerges*. (Evans, 1982, p.153, n.20)

---

<sup>12</sup> Would he be able to distinguish himself from all other things by thinking of himself as 'here'? After all he still has a dispositional connection to a place. No, not if he's whizzing through space.

But whilst this may help in explaining how ‘here’ can refer in the absence of information, it does not help with ‘I’, for persons *are* objects. As O’Brien points out,

it seems unclear how the fact that a subject is disposed to have their thinking controlled by information which may become available could explain how it is that they succeed in identifying themselves when no such information is available (O’Brien, 1995, p.237)

It seems to me that this is just the same problem that Evans faces in accounting for how it is that we entertain non-fundamental ‘I’-Ideas. For one might be tempted to say that the anaesthetized subject manages to entertain ‘I’-thoughts by being able to employ a non-fundamental ‘I’-Idea. That is, although he does not possess his own fundamental ground of difference, he does know what it would be for the propositions ‘ $\delta$  is  $F$ ’ and ‘ $I = \delta$ ’ to be true. And this all seems fine, since there is no suggestion that the anaesthetized subject has forgotten that he is a person. If he had no idea that he was a person, and thus no knowledge as to what it would be for a fundamental identification of himself as a person to be true, then it becomes less plausible to think that he could entertain ‘I’-thoughts anyway.

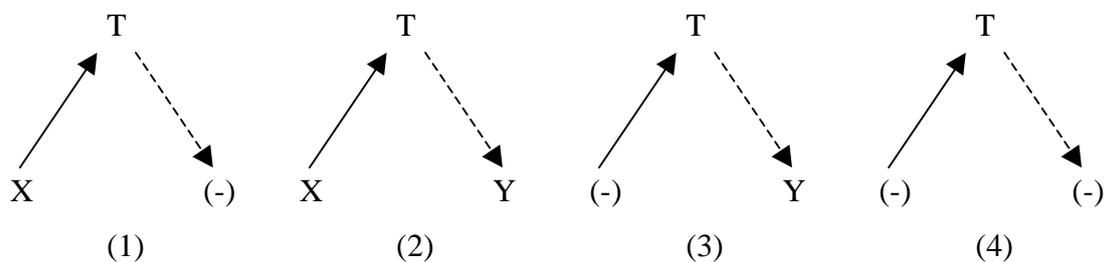
But this just brings us back to the problem of how it is that ‘I’-thoughts which employ non-fundamental Ideas of ourselves can conform to Russell’s Principle. I do not see how Evans can answer this objection and so I think that there is a difficulty at the heart of Evans’ theory of self-identification. The problem is that whilst we want to be able to say that a subject can entertain non-fundamental ‘I’-Ideas, it is difficult to see how such Ideas might conform to Russell’s Principle. This difficulty comes to the fore in the following section’s discussion of the possibility of massive reference failure. Evans’ view that our ‘I’-thoughts might be subject to massive reference failure poses a serious threat to the tight connection between self-consciousness and self-reference that the defender of the Body Claim requires.

### **3.5 The Possibility of Massive Reference Failure**

The reason why I have discussed Evans’ account of ‘I’-Ideas, and the theory of singular thought, at such length is that if the theory is correct it opens up the possibility of a self-conscious subject who fails to satisfy the transcendental conditions of first-person reference. If such a possibility is allowed, the defender of

the Body Claim will be stopped in their tracks, for however convincingly he or she argues that embodied experience is a transcendental condition of self-reference, nothing will follow concerning self-consciousness. This section falls into three parts. First I will outline Evans' views concerning the possibility of reference failure for information-based thoughts in general. I will then illustrate how this impacts on the first-person case, generating the possibility of massive reference failure. Finally I will argue that there are problems with Evans' view. We have already seen that Evans' owes us an account of non-fundamental 'I'-Ideas. This worry looms large in the present section also. The problems, I argue, are generated by Russell's Principle on the one hand, and by Evans' demonstrative model of 'I'-Ideas on the other.

Information-based thoughts are thoughts which must both satisfy Russell's Principle and which require an information-link. According to Evans, there are four different ways in which a subject can essay (in the sense of 'attempt to think') an information-based thought yet fail, two of which correspond to each of these aspects. They are: (1) There is a unique object from which the information derives, but no unique object is identified by the subject's mode of identification, (2) There is a unique object from which the information derives, but a different unique object is identified by the subject's mode of identification, (3) There is no unique object from which the information derives, but there is a unique object identified by the subject's mode of identification, (4) There is no unique object from which the information derives, and there is no unique object identified by the subject's mode of identification. These four cases are represented by Evans as follows (unbroken lines represent derivation of information; broken lines represent purported modes of identification)<sup>13</sup>:




---

<sup>13</sup> See Evans, 1982, p133.

An example of case (1) would be the following<sup>14</sup>: there were two Polish grocers in my home town, but I only remember one of them. The information I retain is derived exclusively from grocer A, yet I cannot distinguish grocer A from grocer B. On Evans' view, I fail to think about grocer A.

An example of case (2) would be the following: I appear to see a cat in front of me. But it is an hallucination, the information of which derives from a cat in another locality. However, there is an exactly similar cat just where I am hallucinating a cat. On Evans' view I fail to think about either cat.

An example of case (3) would be the following: I seem to remember a friend of my parents called Doug. But my apparent memories derive from a dream I once had. However, my parents *did* have a friend called Doug, whom I never met. According to Evans, I fail to have a thought about Doug.

Finally, an example of case (4) would be the following: I hallucinate a pink elephant. The information derives from nowhere, and there is no pink elephant. According to Evans I fail to have a thought about that pink elephant.

Evans' view is that in these situations, whilst I believe myself to be thinking a thought, I have actually failed to do so. That is, I am subject to an illusion of thought<sup>15</sup>. Furthermore, and this is important, precisely similar things can be said of 'I'-thoughts. Evans writes,

It seems to me a corollary of the reflections in this chapter that our ordinary thoughts about ourselves are liable to many different kinds of failings, and that the Cartesian assumption that such thoughts are always guaranteed to have an object cannot be sustained. (Evans, 1982, p.249)

Evans then goes on to list a variety of situations in which, he claims, a subject's essaying of an 'I'-thought would fail. These situations include the following:

- (a) The information that feeds into one's 'I'-Idea via one's special information channels may derive from more than one object (e.g. if one received kinaesthetic

---

<sup>14</sup> This example is from Evans, 1982, p.78.

<sup>15</sup> Evans' view here is radical. To see this, it need only be pointed out that, if substantiated, it undermines the cogito. For one can no longer be certain, and hence it becomes possible to doubt, that one is thinking. However, since Evans is happy to allow that when one has an illusion of thought there is *something* going on in one's mind, we might rewrite the cogito as, 'There is something going on in my mind, therefore I exist'.

information from another person's body as well as one's own) (See Evans, 1982, pp.249-250).

- (b) One's dispositions to act may be manifested in an object other than that from which one's kinaesthetic etc. information derives (See Evans, 1982, p.250).
- (c) One may be a subject with two 'control centres' linked to a single body (See Evans, 1982, p.255, n.74).
- (d) One may be an organism with a 'control centre' outside of one's body, and survive the destruction of one's body (See Evans, 1982, pp.254-5).
- (e) The information that feeds into one's 'I'-Idea may derive from nothing (e.g. if one were a (congenital) brain in a vat) (See Evans, 1982, p.250).
- (f) One may be receiving no information, via one's special information channels, about one's position in space<sup>16</sup>.

These situations represent the possibility of massive reference failure. A subject in any one of these situations would be self-conscious, in that they would believe and subjectively feel themselves to be thinking first-personally. Yet they would not satisfy the conditions of self-reference. It follows that self-consciousness, even in the central cases, does not require self-reference. If this is the case, the defender of the Body Claim will be unable to mount a self-reference based argument for their view.

I want to suggest that there are serious problems with Evans' position here. The culprits, so I shall argue, are his treatment of 'I'-Ideas on the demonstrative (information-based) model, and his adherence to Russell's Principle. However, before I make good on these claims, I must address an objection to the way in which I have placed Evans within the dialectic of the Body Claim. I have represented Evans as supporting the view that self-consciousness is possible in the absence of the transcendental conditions of self-reference. But it seems that *strictly speaking* this is not correct. Evans believes that a subject can attempt to think an 'I'-thought and fail, yet it seem to the subject as if the attempt had succeeded. That is, it can seem to a subject that he is thinking 'I'-thoughts when, in fact, he is not. I have taken this as a challenge to the view that self-consciousness necessarily involves self-reference (for if there is no 'I'-thought there is no self-reference). But one might well object that

---

<sup>16</sup> As should be clear from the previous section, I have added (f) to the list. Evans would not have placed it there, but I think that he should.

Evans' position does not really pose any such challenge. For, as I defined self-consciousness in chapter one, it is the ability to think first-personally; to think 'I'-thoughts. Thus, a subject who failed to do so would not really be self-conscious. If such a subject is not self-conscious, this is no threat to the view that self-consciousness necessarily involves self-reference.

My response to this is to define self-consciousness 'phenomenologically'. Thus, if it seems to a subject that he or she is thinking 'I'-thoughts, that is sufficient for him or her to be considered self-conscious. Indeed, I think that this is the intuitive view. 'From the inside' such a subject may be entirely indistinguishable from a 'properly' self-conscious subject. Any definition which admitted that one was self-conscious, and the other not would, I suggest, be deficient. Thus, it should be clear that the definition of self-consciousness given in chapter one is not quite correct. Self-conscious subjects are those who are in a position such that it seems to them that they can think first-personally. Given this definition of self-consciousness, what we might call phenomenological self-consciousness, Evans' position does indeed pose a threat to the claim that self-consciousness necessarily involves self-reference. Having made this qualification to the definition of self-consciousness, I will revert to the original turn of phrase, and continue to speak of self-consciousness as first-person thought; as a property of thoughts.

With this objection out of the way we can take a closer look at the situations in which Evans thinks an essayed 'I'-thought would fail. To begin with, I will assume that in all of the situations (a)-(f) there is no query as to which body is the subject's body. That is, there is exactly one body which is the subject's body. Specifically, in situation (a) the subject is receiving kinaesthetic information from a body *which is not his or her own*. And, in situation (b), the subject's dispositions to act are manifested in a body which is *not the subject's own body*. In situation (c), I take it that the suggestion is that there are *two subjects* here, both of whom 'share' a body. The first question we must answer is how (a)-(f) relate to (1)-(4). That is, how the cases given by Evans of failed 'I'-thoughts, map on to his schematic account of the possibility of ill-grounded information-based thoughts.

It is not immediately clear how these situations relate to the four different kinds of ill-groundedness. Situations (a) and (b) are, I suggest, instances of either case (3) or case (4). In each situation we are told that there is no unique object that is the

source of the subject's information<sup>17</sup>. Thus they are both either instances of case (3) or of case (4). However, we are not told of the subject's mode of identification, so we are unable to decide between the two cases.

It seems to me that, as described, situations (c), (d), (e) and (f) are all either instances of case (1) or of case (4). The difficulty in situation (c) is that there are two subjects here and the mode of self-identification they employ is not uniquely identifying. Each subject's mode of identification nets two objects (the two subjects). Thus, we have an instance of either case (1) or case (4). Since we are not given any details concerning the subject's information links, we are not in a position to decide between the two.

Situations (d), (e) and (f) are problematic not because of a lack of uniqueness, but because the subject's mode of identification fails to net an object at all. These subjects do not have uniquely identifying conceptions of themselves, as they are unable to locate themselves in space. Again, we do not have enough details concerning the subject's information-links to be in a position to decide whether these situations are instances of case (1) or of case (4).

To complete the list, an instance of case (2) would be one in which a subject received information (or misinformation) from him or herself, yet which failed to identify him or herself and which, through chance, correctly identified another subject.

Evans' claims that a subject in one of these situations essaying a first-person thought would fail. That is, a self-conscious subject would fail in his or her attempts to think first-person thoughts, and thus fail to self-refer. Hence, if Evans is correct, the Body Claim is in trouble since self-consciousness does not necessarily involve the capacity for self-reference.

These cases can be divided into those in which something goes wrong on the information-link side and those in which something goes wrong on the mode of identification side<sup>18</sup>. In the remainder of the section I will argue for the following claims. First, those cases in which something goes wrong on the information-link side are not intuitively described as instances of reference failure. Evans' description of

---

<sup>17</sup> It is easy to be misled into treating (b) differently from (a) due to its involving both input and output. This would be a mistake. Output is on the information-link side rather than the mode of identification side.

<sup>18</sup> And those in which something goes wrong on both sides. I shall not discuss such cases separately.

them as such is forced by his treating 'I'-Ideas on the demonstrative model. If we abandon this model we enable ourselves to say what seems intuitively correct about these cases. Second, those cases in which something goes wrong on the mode of identification side are only correctly described as instances of reference failure if we accept Russell's Principle. It is tempting to describe such cases as perfectly acceptable, and self-referential, instances of subjects employing non-fundamental 'I'-Ideas. What prevents us from saying this is Russell's Principle. If we were to reject Russell's Principle we could reject Evans' view that these are cases in which the subject does not meet the requirements of self-reference. Hence, if we reject both the demonstrative account of 'I'-Ideas and Russell's Principle, we can rebut the possibility of massive reference failure, thus reinstating the tight connection between self-consciousness and self-reference.

I shall begin with those cases in which something goes wrong on the information-link side. Case (a) is the simplest of these. Here the problem is that the information that feeds into one's 'I'-Idea derives from more than one object. An analogous case for 'this' supports the view that there is no reference here. "Suppose one uses the expression 'this cup' when one is seeing a cup and feeling a cup (in fact, though one does not know it, *two* cups)" (Evans, 1982, p.250, n.67). In this situation, it is intuitive to say that the use of the expression 'this cup' has failed to refer. But 'I' is supposed to follow the model of 'here' closer than the model of 'this'. So, we should see whether we can construct an analogous case for 'here'. Suppose one is standing in a quiet room, but one's perceptual (visual and auditory) information derives from some other (noisy) place, and one utters 'It is loud here'. In this situation we are less inclined to think that reference has failed, and more inclined to think that one has uttered something false. The reason for this is that there is a fact of the matter about which place the subject is at. This gives us a reason to privilege one place over the other as the reference of 'here'. But now we can see that exactly the same point applies to situation (a). In the envisioned situation, one's kinaesthetic information derives from someone else's body, as well as one's own. But, as I said above, there is no question as to which body is one's own (this is built into the description of the case). So we have a reason to privilege one person over the other as the referent of one's 'I'-thought. Thus, it is plausible to say that one's thought is (presumably) false rather than a failure of reference.

Situation (b) is more complex, but essentially the same point applies. Whereas (a) involves input, (b) involves output. Situation (b) is seriously under-described: we do not know whether the subject is conscious of his or her dispositions to act being manifest in this other body. If so, there must be some deviant input as well. If not, then it is hard to see exactly how it is supposed to be problematic. For, suppose Bruce suffers from entire-body paralysis. We have yet to be given any reason to think that he is incapable of referring to himself with ‘I’<sup>19</sup>. What difference would it make that his dispositions to act are manifested in another body? Presumably none. We have created a deviant information-channel here but, as before, we have every reason to privilege the information channel which leads to the subject’s actual body. So, we should assume that Bruce gets some feedback from this other body. And now the question must be whether there is any reason to think that Bruce’s attempts to refer to himself with ‘I’ would fail. Not intuitively, since situation (b) is now similar to situation (a), and it is not intuitively plausible to describe situation (a) as a failure of reference.

The point that I have been pressing is that it is wrong to view ‘I’-thoughts as information-based, i.e. on the model of demonstrative thought. As Evans’s himself notes, whilst demonstrative thoughts are “information-based thoughts *par excellence*, [‘I’-thoughts] do not seem to depend necessarily either upon the subject’s actual possession of information...or upon the actual existence of an information-link” (Evans, 1982, p.152)<sup>20</sup>. But treating ‘I’-thoughts as information-based at all means that we have to say some highly implausible things concerning the cases just described. There are two aspects to information-based thoughts. First, such thoughts are governed by their relevant information channels. Second, those information channels are essential to those thoughts. Whilst it seems reasonable to think that the first aspect correctly applies to ‘I’-thoughts, it is not plausible to think that the second does. When the information channels go wrong, the intuitive view is that the subject’s capacity for ‘I’-thoughts remain unaffected. This is why treating ‘I’-thoughts as demonstrative, information-based, thoughts is so counterintuitive. If we depart from Evans on this matter, as I believe we should, cases (a) and (b) no will longer be treated as cases of massive reference failure.

---

<sup>19</sup> But see my discussion of the action argument in chapter eight.

<sup>20</sup> Evans is actually talking about ‘here’-thoughts at this point, but I presume that he would say the same of ‘I’-thoughts.

Let us move on to the situations in which something goes wrong on the mode of identification side. First, take situation (c). On the (questionable) assumption that this is a conceivable situation, it is at first difficult to see why there should be any failure of first-person thought. For there are two subjects here and each subject's 'I'-thoughts will refer to that subject. But I take it that the problem, as noted above, is that neither of these subject's modes of self-identification will be uniquely identifying. Therefore, neither subject is in a position to distinguish themselves from all other things. Thus, by Russell's Principle, we must deny first-person thought to these subjects. In the next section I shall argue that Russell's Principle should be rejected, and therefore that this situation is not problematic in the way Evans claims.

Situation (d) introduces rather more complications. Concerning situation (d), Evans asks us to imagine

organisms whose control centre is outside the body, and connected to it by communication links capable of spanning a considerable distance. An organism of this kind could have an idea of itself like our own, but if it did, it would be unable to cope with the situation that would arise when the control centre survived the destruction of the body it controlled. Thinking like us, the subject would of course have to regard itself as somewhere, but in this case it would not make any sense to identify a particular place in the world as the place it thought of as *here*...Because its 'here' picks out no place, there is no bit of matter, no persisting thing, which the subject's Idea of itself permits us to regard as what it identifies as itself. Here, then, we have a very clear situation in which a subject of thought could not think of itself as 'I'; its 'I'—its habitual mode of thought about itself—is simply inadequate to the situation. (Evans, 1982, pp.254-5).

This passage may seem rather strange. It seems clear that this organism (let's call her Shelia) does not have a fundamental Idea of herself, for she does not know her spatial location. But mightn't Shelia have a non-fundamental Idea of herself? For possession of a non-fundamental Idea of an object does not require one to be in possession of that object's fundamental ground of difference. To have a non-fundamental Idea of a (physical) object it is sufficient that one have a ability to conceive of the object of one's thought as located in such-and-such a place. But Evans provides us with no reason to deny that Shelia has this ability. And if this is the case, then there seems to be no reason to think that Shelia cannot perfectly well think 'I wonder where I am', referring to herself in the process.

As I have already said, it is hard to see how non-fundamental 'I'-Ideas conform to Russell's Principle. And it does look as though Shelia's thought would fail to so conform. This, no doubt, is the source of Evans' denial of Shelia's ability to think of herself as 'I'. But here we just have a vivid example of the tension between Russell's Principle and the notion of a non-fundamental 'I'-Idea. I shall argue below, that it is Russell's Principle that has to go.

Situation (e) can be thought of as something like a combination of (a) and (d). In situation (a), the subject's information upon which physical self-ascriptions (but not his mental self-ascriptions) are based derives from another's body, here it derives from nothing. An analogous demonstrative case would be something like this: If I hallucinate a tennis ball, I cannot refer to anything using the demonstrative 'this ball'. For there is no ball. But the analogy is not very close. Admittedly, for the brain in the vat, there is no body. But it is not the case that there is no subject. For, whilst it is true to say that a demonstrative (or any other expression, for that matter) will fail to refer if its intended object does not exist, this does not correctly describe case (e).

I take it that the real problem that Evans finds with (e) is more like the problem we discerned with (d). That is, the brain in the vat does not have a fundamental Idea of itself, for it does not know its current location; all the information it has about its current location is illusory. But, if this is the problem, we can ask the same question as we asked about (d): mightn't the brain in the vat have a non-fundamental Idea of itself? That is, it knows what it would be for an arbitrary proposition of the form ' $\delta = a$ ' to be true<sup>21</sup>. What could justify refusing to attribute this kind of Idea of itself to the brain in the vat? Presumably the claim that it would violate Russell's Principle.

Finally, we come to case (f). It has already been admitted by Evans that it is intuitively plausible to view the subject of complete sensory deprivation as capable of first-person thought. I suggested above that Evans can allow first-person reference to the subject in sensory deprivation if he allows that the subject has a non-fundamental Idea of himself. But, again, this raises the difficulty of reconciling non-fundamental 'I'-Ideas with Russell's Principle. We can see, then, that the situations described by Evans as involving something going wrong with the mode of self-identification are plausibly described as involving non-fundamental 'I'-Ideas. If this claim could be

---

<sup>21</sup> Would ' $\delta$ ' range over fundamental Ideas of persons or brains?.

made out, then such cases would not constitute examples of massive reference failure. Of course, the problem with the suggestion is that it is entirely unclear how non-fundamental 'I'-Ideas conform to Russell's Principle. If we adhere to Russell's Principle, we seem forced to admit that there is reference failure here. It seems that the time has come to assess Russell's Principle itself.

### 3.6 Against Russell's Principle

If Russell's Principle is rejected we will have no reason to think that the subjects described in the previous section are incapable of self-reference. This will save the Body Claim from an early bath. In this section I argue against Russell's Principle<sup>22</sup>. To begin with it must be pointed out that Evans accepts Russell's Principle only at the level of thought, not at the level of language. Evans writes that,

The abandonment of the principle of identification [Russell's Principle] at the level of saying is a trivial consequence of the distinction between what one says and what one intends to express. Its abandonment at the level of *belief* or *thought* would be an extremely significant move. What has happened is that the former has been mistaken for the latter. (Evans, 1982, p.76, n.18).

This is important as it opens up a way of criticising Evans' position directly. For it would appear that any of the following views, in descending order of philosophical sophistication, would render Evans' position problematic: (i) The conditions of linguistic reference are identical to the conditions of reference in thought, (ii) The only way to analyse thought is via the analysis of language, (iii) That which I can speak about, I can think about. Of course, if one of these claims could be

---

<sup>22</sup> It is worth noting that rejecting Russell's Principle is not the only way to save the connection between self-consciousness and self-reference from Evans' attack. One could argue that subjects in the situations described in the previous sections are able to identify themselves descriptively as 'the thinker of this thought', or 'the subject of these conscious states'. If this were possible, one could reject the possibility of massive reference failure, whilst remaining neutral on the question of Russell's Principle. This strategy is well represented by Peacocke, 1983, Ch.6. Whilst I am not at all convinced that this isn't a perfectly adequate response to Evans, I do not follow that route for a number of reasons. First, it is extremely controversial whether one really can identify oneself descriptively in this way. For one thing it is not absolutely clear that we can demonstrate our thoughts *at all*, and even if we can there is a highly influential line of thought, due to Strawson, 1959, p.97, to the effect that the demonstrative identification of a thought is parasitic upon the demonstrative identification of the person whose thought it is. If this is correct, the proposed account would be circular. My second reason for pursuing the route of rejecting Russell's Principle is that it is useful for the discussion in the following chapters. Having rejected Russell's Principle here, I go on to make further use of that rejection.

substantiated, the proponent of Russell's Principle could move in either direction. That is, they could reject Russell's Principle at the level of thought, or they could adopt it at the level of language.

I will not be arguing against Russell's Principle in this way. Rather, the case that I will put against Russell's Principle is indirect, and will come in two parts. First, I argue that the acceptance of Russell's Principle has some extremely counterintuitive consequences concerning what one is capable of thinking about. Second, I will undermine Evans' own theoretical arguments for the principle. Together these considerations lead us to a rejection of the principle, since a principle with counterintuitive consequences surely requires a strong theoretical backing, and this is something that Russell's Principle lacks. In an appendix to *The Varieties of Reference* Evans claims that the rejection of Russell's Principle amounts to the rejection of the Fregean notion of sense. At the end of the section I will argue that this need not be the case, and will sketch out some positions that could be taken by one who denies the principle.

That Russell's Principle has counterintuitive consequences is something which should have become clear in the preceding discussion of reference failure. For instance, it is a consequence of Russell's Principle that a subject in the position of Shelia cannot think to herself 'I wish my body hadn't just been destroyed'. For Shelia is unable to distinguish herself from all other things; she cannot locate herself spatially. But it is not only concerning 'I'-thoughts that Russell's Principle has unwanted consequences. Much the same can be said for both demonstratives, and proper names.

Suppose a subject sees an apple in a mirror at a fairground stand<sup>23</sup>. The mirror is part of a larger set-up of mirrors, the result of which is that the subject cannot locate the apple in space. If Russell's Principle is true, the subject cannot think the demonstrative thought 'I wonder where that apple is', for the subject cannot distinguish the apple from all other things by locating it<sup>24</sup>. In a similar vein, since audition does not provide the requisite identification information, I cannot entertain a

---

<sup>23</sup> This example is from Peacocke, 1983, pp.170-1. Campbell, 2002, pp.111-112, offers a similar example.

<sup>24</sup> This may not be an ideal example, for the subject may still have the *ability* to locate the apple. See Peacocke, 1991, pp.124-125. For the claim that it may be indeterminate whether a subject has such an ability, see Millikan, 2000, Ch.13.

demonstrative thought about the drill outside my room on the basis of hearing it<sup>25</sup>. Again, suppose a subject is put inside a wooden box and taken on a long journey. If Russell's Principle is true, then that subject cannot think 'I wonder what the weather's like here, wherever here is'. For, once again, he cannot distinguish the place where he is from all other places.

In the case of names, adherence to Russell's Principle leads to many counterintuitive consequences. Due to my inexcusable lack of knowledge concerning the history of philosophy I cannot distinguish between Fichte and Schelling. All I know about either of these two men is that they were post-Kantian German philosophers. So, since I cannot distinguish Fichte from all other things it follows that I cannot entertain thoughts about Fichte, since I cannot distinguish him from all other post-Kantian German philosophers<sup>26</sup>. This, I suggest, is more than a little surprising.

So, Russell's Principle has consequences that strike us as counterintuitive. Some commentators have felt that simply pointing this out is all that one needs to do in order to reject the principle<sup>27</sup>. But this surely is not enough. For Evans provides arguments for Russell's Principle, and if the principle is to be rejected these arguments must be shown to be unconvincing. Evans deploys (at least) two distinct arguments for Russell's Principle. I shall argue that neither of these are compelling<sup>28</sup>.

Evans' first argument utilises the Generality Constraint. Evans claims (1982, §3.3) that certain supposed counterexamples to Russell's Principle violate the Generality Constraint and so are unacceptable. The first supposed counterexample is of the same kind as my 'Fichte' example above. We are asked to suppose a child to have been "introduced to the name 'Socrates' by hearing simply that Socrates was a Greek philosopher" (Evans, 1982, p.73). Now,

If the ignorant child has got hold of the widely disseminated piece of information (or misinformation) 'Socrates was snub-nosed', we might well be inclined to say that the child has a true or false belief about Socrates...But the inclination to say that the child has, and is expressing, a belief about Socrates is far less strong when we envisage the child not merely repeating the widely disseminated piece of

---

<sup>25</sup> See Campbell, 2002, p.112-113.

<sup>26</sup> For simplicity I ignore the (real) problem concerning descriptions such as 'the post-Kantian German philosopher who was called "Fichte"'.

<sup>27</sup> See, for example, Bach, 1987, pp.41-5.

<sup>28</sup> My criticisms of Evans' arguments for Russell's Principle are greatly influenced by Rozemond, 1992-3, and Sainsbury, 1985. Also see Loar, 1984.

information, but uttering the words 'Socrates is fat' (say), perhaps as a result of some confusion. (Evans, 1982, p.75)

The question is whether the child is in a position to entertain the thought, of Socrates, that he is fat. If the answer to this is no, then the supposed counterexample to Russell's Principle violates the Generality Constraint. We can, for the sake of argument, agree with Evans that anything which violates the Generality Constraint must be rejected. But I see no reason to think that this example does so.

In what way is the child confused? Let us suppose that she wrongly supposes all philosophers to be fat<sup>29</sup>. She then, perfectly naturally draws the conclusion that Socrates is fat. Here there is no temptation to think that such a child is not in a position to entertain the thought that Socrates is fat. In a similar vein, I have no distinguishing knowledge of Fichte. It would, then, be a counterexample to Russell's Principle if it were correct to ascribe to me the belief that Fichte was a post-Kantian German philosopher. According to Evans' present argument, my supposed belief would violate the Generality Constraint, since I cannot be credited with the capacity to entertain the thought that Fichte loved his mother. But the claim that I cannot entertain such a thought is simply offered with little by way of justification. So, I think that Evans' first defence of Russell's Principle, using the Generality Constraint, is unconvincing.

Evans' second argument for Russell's Principle involves the example of two identical steel balls<sup>30</sup>. Suppose that a

subject briefly sees one ball rotating by itself on one day, and the other on a later day. And let us further suppose that the subject retains no memory of the first episode...Suppose, finally, that many years later our subject reminisces about 'that shiny ball' he saw many years earlier. If asked which ball he is thinking about, our subject cannot produce any facts which would discriminate between the two. (Evans, 1982, p.90)

If it is correct to say that the subject is thinking about the second ball, then we have here a counterexample to Russell's Principle. But Evans does not think that it is correct to so describe the subject. The reason for this is that it would be wrong to

---

<sup>29</sup> This line of thought is filled out in Rozemond, 1992-3, §II.

<sup>30</sup> The steel balls example is analogous to the Polish grocers example mentioned in the previous section.

ascribe to this subject an Idea of the ball. For an Idea of an object is an ability to think about that object, but there is no ability that this subject has which would make his Idea an Idea of the second ball rather than the first,

There is no question of his recognising the ball; and there is nothing else he can do which will show that his thought is really about one of the two balls (about *that* ball), rather than about the other. (Evans, 1982, p.115)

Therefore, the subject does not have an Idea of the ball, and so cannot entertain thoughts about it. There are two ways in which I want to object to this. First, I simply deny that there is no ability that the subject has, which would make his Idea an Idea of the first rather than the second ball. For the subject has the ability to remember, reason and deliberate about the second ball, but lacks the ability to remember, reason and deliberate about the first<sup>31</sup>. Thus, our subject can be distinguished from a subject who could think about the first, but not the second ball. Does this beg the question against Evans? I don't think so. Evans challenges his opponent to say what it is that the subject can do which makes his thought a thought about one ball rather than another. The subject's capacity to remember the ball, and so on, constitutes his Idea of the ball. True, *we* cannot distinguish the subject from another subject thinking about the other ball. But why should this matter? The capacities of the two subjects are different, since these capacities must be individuated with reference to their objects.

The second objection is slightly different. It involves the thought that if Evans is right concerning the steel ball case, then this will lead to some extremely counterintuitive results. For suppose that, rather than see two balls and forget about one, the subject sees only one ball, since the other ball was hidden behind a curtain<sup>32</sup>. Now, if Evans is right to say that there is no ability that the first subject had that could show that he was thinking about the second rather than the first ball, then there can be no such ability in this new variation of the example. For, "There is no question of his recognising the ball; and there is nothing else he can do which will show that his thought is really about one of the two balls (about *that* ball), rather than about the

---

<sup>31</sup> See Rozemond, 1992-3, p.288, Peacocke, 1983, p.174, and Loar, 1984, p.50. McDowell, 1990, p.260, claims that this objection misses the point but I don't see how.

<sup>32</sup> See Sainsbury, 1985, p.133-4.

other” (Evans, 1982, p.115)<sup>33</sup>. But this conclusion is absurd. Thus, we should not be convinced by Evans’ second defence of Russell’s Principle. Since Russell’s Principle has such counterintuitive consequences, and Evans’ theoretical defence of the principle is unconvincing, we should not accept it.

Before we leave the subject of Russell’s Principle it is necessary to address one last issue. In an appendix (Evans, 1982, pp.80-85), Evans suggests that the rejection of Russell’s Principle is tantamount the rejection of the notion of sense for singular thoughts. If this were the case it would be disastrous to anyone who thinks that there are singular thoughts and is convinced by Frege’s Puzzle. To support his suggestion, Evans points out that, “it is quite obscure how, if one mental state represents a particular object in virtue of one sort of causal relation to it, and another mental state (of the same subject) represents that object in virtue of another sort of causal relation to it, the sheer difference between causal relations could generate a difference in *content* between the two mental states, given that it need not in any way impinge on the subject’s awareness.” (Evans, 1982, p.83). Evans’ point seems to be something like the following: Russell’s Principle provides an epistemic constraint on the ascription of thought content. The only serious alternative to a theory of singular thought adhering to Russell’s Principle is a causal account. But a causal account can have no place for sense, as sense is a *cognitive* notion, it is whatever it is that explains the cognitive significance of language and thought. The rejection of Russell’s Principle would allow, to borrow Wettstein’s useful phrase, ‘semantic action at epistemic distance’ and this is entirely mysterious<sup>34</sup>.

Evans overstates his case here. Indeed, I submit that there are at least three, broadly outlined, positions that can be occupied by those who deny Russell’s Principle:

(A) One could reject Russell’s Principle yet continue to maintain that there are non-trivial epistemic conditions on reference. I take it that the view proposed by Peacocke, 1983, Ch.7, falls into this category. The point is that one can continue to accept that

---

<sup>33</sup> One might be tempted to say that the ball can be distinguished as ‘the one that I saw’, but Evans himself rules out such a move on the grounds that “thinkers...will not in general resort to this way of identifying the object of their thought” (Evans, 1982, p.117). Indeed, it seems that if this is all that is required, there would never be any counterexamples to Russell’s Principle, for a subject could always distinguish the object of his thought as ‘the object causally responsible for my current thought’.

<sup>34</sup> Wettstein, 1991, Ch.11. Some discussion of the issues surrounding Russell’s Principle can also be found in Chs.8, 9 and 10 of that collection.

reference requires something which can correctly be described as ‘knowing what one is thinking of’, but maintain that such knowledge need not amount to anything so strong as Evans’ discriminating knowledge<sup>35</sup>.

(B) One could reject all epistemic conditions on reference, allowing semantic action at epistemic distance, but continue to insist that Frege’s Puzzle can be solved by a causal account. For example, it might be argued that not only what one is thinking about, but also how one is thinking about it is determined causally. Thus, pace Evans, a difference in the causal ancestry of two thoughts about *a* will (sometimes) result in a difference in the way that one can think of *a*, and hence in a difference in content.

(C) One could reject Russell’s Principle, allow semantic action at epistemic distance and furthermore deny that Frege’s Puzzle needs to be solved. This is the position adopted by Wettstein, 1991, Ch.11.

To my mind the first option is the most attractive, and whilst I do not pretend to have such an account readily available, I have some confidence that the notion of sense can be saved by producing an epistemic condition on reference which falls well short of Russell’s Principle. My current point is simply that Evans has surely overstated the case for Russell’s Principle in claiming that its rejection amounts to the rejection of sense.

Given the rejection of Russell’s Principle, we have no problem in allowing subjects to entertain non-fundamental ‘I’-Ideas. This allows us to account for all the supposed cases of reference failure mentioned in the previous section, and so the challenge posed by Evans’ account of self-identification to the tight conceptual connection between self-consciousness and self-reference can be put aside. Evans has not managed to show that there could be a self-conscious subject who failed to satisfy whatever conditions there are on self-reference.

---

<sup>35</sup> In chapter six I offer an extremely weak epistemic condition on reference. This has nothing to do with the notion of sense however.

## 4. Embodied Experience

### 4.1 Bodily-Awareness

The Body Claim links self-consciousness and embodied experience. In the preceding chapters I have said a great deal about self-consciousness but very little about embodied experience. Common wisdom says that we have five senses; sight, hearing, smell, taste and touch. These senses provide us with a means of gaining information concerning the world around us. But in addition to these we possess certain means of gaining information about our own bodies. These include our awareness of the position, orientation, movement and size of our limbs (proprioception and kinaesthesia), our sense of balance, and our awareness of bodily sensations such as pains, tickles and sensations of pressure and temperature. We can group these together under the heading ‘bodily-awareness’. In this section I will argue briefly for three claims concerning bodily-awareness. First, bodily-awareness is an awareness of one’s own body. Second, bodily sensations are experienced as located in or on one’s body. Third, in bodily-awareness one is aware of one’s own body *as* one’s own body.

The first of these claims, may seem to be a truism, but I take it to be controversial in two respects. First, there are philosophers who would argue that bodily-awareness cannot be treated as a unitary phenomenon at all, and second there are those who would argue that bodily-awareness cannot be treated as a form of awareness of an objective part of the world. In support of the first of these objections it might be pointed out that the various phenomena that I have picked out as components of bodily-awareness are disparate in nature, and are subserved by a multitude of information processing systems. Furthermore, there is no organ of bodily-awareness. What sense is there, then, to speak of a unified phenomenon of bodily-awareness? This objection is easily met, however. For bodily-awareness is unified phenomenologically. The various ways of gaining information of our own bodies that I have grouped together are unified in their providing the subject with an awareness of the body ‘from the inside’<sup>1</sup>. I can be aware of my body through, for instance, vision but this awareness of the body is from the outside. The phenomena

---

<sup>1</sup> See Martin, 1993.

grouped together as bodily-awareness, on the other hand, do not provide an awareness of the body from a perspective external to the body itself. Thus, we are justified in speaking of the unified phenomenon of bodily-awareness.

The second objection is that bodily-awareness does not constitute an awareness of the body<sup>2</sup>. Philosophers who have made this claim have tended to concentrate on bodily sensations. On this view bodily-sensations are subjective states which are located entirely within the mind. Bodily sensations do not represent, and so provide the subject with no access to, the objective world<sup>3</sup>. Further, the proponent of this view is likely to maintain that our judgements of our own bodily position and orientation are not based on an awareness of the body; that we have ‘knowledge without observation’ of the position and orientation of our limbs<sup>4</sup>. This view is sometimes supported by the observation that we do not judge the position and orientation of our bodies via an awareness of bodily-sensations<sup>5</sup>. But it should be pointed out that even if this last observation were true, it would still be consistent with the view that proprioception is a non-sensational, or not purely sensational, awareness of the body.

This objection to the view that bodily-awareness is an awareness of the body is not convincing. As has been argued by several philosophers, there are good reasons for thinking that bodily-awareness involves an awareness of the objective body<sup>6</sup>. The most compelling reason to adopt this view is that, as O’Shaughnessy, 1980, has made clear, bodily sensations are presented as located within one’s body. When one has a sensation of pain, there is an answer to the question of where it is that one feels it. Furthermore, that answer can only be given by stating a location in or on one’s felt body. This lends plausibility to the view that bodily-awareness is an awareness of one’s body as being, say, painful or hot etc. As Armstrong points out, adopting the view of bodily-awareness as an awareness of the body, allows us to say that

The locations of such sensations in the body is therefore an *intentional* location. To say that I have a sensation of pressure in the small of the back is to say that physical pressure *seems* to be occurring there. (Armstrong, 1968, p.310).

---

<sup>2</sup> Some would allow that bodily-awareness is an awareness of the body but maintain that it is not a *perceptual* awareness. I remain neutral on this issue.

<sup>3</sup> See McGinn, 1982, pp.8-9.

<sup>4</sup> This claim is particularly associated with Anscombe, 1957 & 1962.

<sup>5</sup> See Wittgenstein, 1953, Pt.II, §viii.

<sup>6</sup> See Armstrong, 1962 & 1968, Ch.14; O’Shaughnessy, 1980, Vol.I, Ch.5; and Martin, 1993 & 1995.

The view of bodily-awareness as purely subjective has a far more difficult time accounting for this felt location of sensation. If bodily sensations are purely subjective, qualitative states with no intentional content, it is difficult to see how one arrives at the bodily location of the sensation. The most natural suggestion seems to be that each bodily location is associated with a distinctive qualia, or qualitative feel. But this seems incredible. For if the location of sensation is to be accounted for in this way it rules out the possibility of one's having phenomenologically identical sensations in two distinct bodily locations. But on having one's ears pierced, say, it would seem that the pain in each ear is exactly the same, differing only in felt location<sup>7</sup>.

If this is right, we should admit that bodily sensations have spatial content. When one has a bodily sensation one feels as if something is happening in a particular part of one's body. Once this is admitted there is surely no objection to the claim that proprioception and kinaesthesia have spatial content also. Thus, the first two claims to be made out in this section are justified: first that bodily-awareness is an awareness of one's own body, second that bodily sensations have a felt location. The third of the claims to be made in this section, that in bodily awareness one is presented with one's body *as* one's body, is closely tied up with the first two. The idea is that bodily sensations are not simply presented as located in some body or other. Rather they are presented as located in *one's own body*. If I feel a pain in a neck, I thereby feel that I have a pain in *my* neck<sup>8</sup>. Through bodily-awareness whatever I am aware of appears to me to be (a part of) my own body. Martin understands this claim as the following,

Any region in which it seems to one that one could now be feeling sensation will thereby feel to one to fall within one of one's boundaries (Martin, 1995, p.271).

Thus, bodily-awareness is an awareness of one's body which comes hand in hand with a 'sense of ownership'. When I am aware of a body through bodily-awareness that body thereby feels to be my body. This is another way of saying that through bodily-awareness I am presented with my body *as* my body. This, then, is embodied experience. We, each of us, have an awareness of our own bodies through which we

---

<sup>7</sup> See Martin, 1993, p.208.

<sup>8</sup> This claim does not conflict with Wittgenstein's claim that "It is conceivable that I feel pain in a tooth in another man's mouth" (Wittgenstein, 1958, p.49), so long as the tooth feels to be one's own.

are aware of sensations as located in or on what is presented as one's own body. This is the core sense that each of us has of his or her own embodiment<sup>9</sup>.

## 4.2 Bodily Self-Awareness

I have outlined an account of bodily-awareness, but it might be thought that bodily-awareness is more than simply an awareness of one's body. Bodily-awareness might be thought of as a form of self-awareness; as bodily self-awareness. In the introduction and chapter one I distinguished between two conceptions of self-consciousness: self-awareness and first-person thought. I also pointed out (see *fig.1* in the introduction) that bodily self-awareness can be thought of as both a form of self-consciousness and a form of embodied experience. The aim of the present chapter is to evaluate that version of the Body Claim which argues that bodily self-awareness is a transcendental condition of first-person thought. The most prominent defender of this version of the Body Claim is Cassam. In *Self and World*, Cassam distinguishes between what he calls 'three grades of apparent presence in the world',

The first grade involves intuitive awareness of oneself as present in the world only geometrically, as something bodiless. The second grade involves awareness of oneself as a point of view which 'has' a body...The third grade of presence involves intuitive awareness of oneself *qua* subject as a bodily presence in the world (Cassam, 1997, p.58)

This is a useful distinction. In the current terminology, we can think of Cassam's 'third grade of apparent presence' as being just the same as bodily self-awareness. Thus, the claim under investigation is whether this third grade of presence is a transcendental condition of first-person thinking. Cassam argues at length that it is. I shall be arguing against Cassam on this issue. Bodily self-awareness is not a transcendental condition of self-consciousness. Furthermore, bodily self-awareness is not something which ordinary human beings enjoy. Since we are self-conscious, in the relevant sense, it could not possibly be a transcendental condition of self-consciousness. Whether the second grade of apparent presence in the world, that one experience oneself as 'having' a body, is a transcendental condition of self-consciousness is a question to which I shall turn in the subsequent chapters. In the

---

<sup>9</sup> Of course there are countless other fascinating aspects to embodied experience not discussed here.

present section I give a rough understanding of what bodily self-awareness is supposed to be. In the following sections I argue first against the view that bodily self-awareness is required for self-consciousness, then against the view that there is such thing as bodily self-awareness at all.

The idea under consideration is that bodily-awareness, the awareness that we have of our bodies ‘from the inside’, constitutes a bodily form of self-awareness. This claim has been explicitly endorsed by Ayers, 1991, Brewer, 1995, Cassam, 1995a, & 1997, Bermúdez, 1998, and de Gaynesford 2002<sup>10</sup>. Hence, we find Ayers saying that

the physical or material self is...neither simply...the postulated subject of experience, nor simply...an ever-present object of experience, but...the *presented* subject of experience and action (Ayers, 1991, Vol. 2. p.286)

The claim here is not just that the body and the self are identical so, in being aware of one’s body, one is aware of oneself. For this would be to obscure the claim that self-awareness is a special form of awareness through which one can be aware only of oneself. The claim is that we have a special *introspective* awareness of ourselves as bodily beings. As Cassam puts the point,

introspective self-awareness is not just awareness or consciousness of what is in fact the subject of one’s thoughts and perceptions...Rather, introspective self-awareness must be understood as awareness of oneself ‘*qua* subject’ (Cassam, 1997, p.4)

This description of self-awareness as introspective helps make clear the sense in which Cassam et al are in direct opposition to Hume’s famous scepticism concerning self-awareness (quoted in chapter one). Hume denies that there is such a thing as introspective self-awareness, Cassam maintains that there is and that it is constituted by bodily-awareness. This strong grade of apparent presence in the world is an experience of oneself not only as embodied, but *as a body*. Bodily-awareness, on this view, is not just an awareness of one’s body, it is an awareness of one’s body as an object which presents itself as oneself.

---

<sup>10</sup> The idea is, arguably, also present in Strawson, 1966a, and Evans, 1982, Ch. 7. Compare also Sutton Morris’ claim that, “Hume, looking into himself, is unable to discover any continuing datum. He overlooks the most obvious candidate, one’s continuing sense of being an embodied experiencer” (Sutton Morris, 1982, p.217). Finally, compare Freud’s well known claim that, “The ego is first and foremost a bodily ego” (Freud, 1923, p.364).

### 4.3 Cassam's Identity Argument

Cassam, 1997, puts forward several arguments for the claim that bodily self-awareness is a transcendental condition of self-consciousness but, in the end, the only one he finds convincing is what he calls 'the intuition version of the Identity Argument'. Cassam's argument is a transcendental argument. In the introduction I made the point that such arguments face a difficulty concerning the controversial relationship between conceivability and possibility (or inconceivability and necessity). I suggested that transcendental claims should be thought of as of the form 'y without x is inconceivable' rather than of the form 'x is a necessary condition of y'. Taking this on board, we should think of Cassam's identity argument as concluding that self-consciousness without embodied experience is inconceivable. I also said that I would mark this point by speaking of transcendental conditions. This is fine, but there is another well known worry concerning transcendental arguments that has not yet been mentioned. Transcendental arguments take y to be some feature of experience that is beyond sceptical doubt (in our case, the existence of self-consciousness). As such, they have often been used as a way of countering scepticism about condition x, for if y is beyond sceptical doubt and is inconceivable without x then it would appear that scepticism about x is inconceivable. But the validity of using transcendental arguments against the sceptic in this way has come in for a great deal of criticism. In particular, Stroud, 1968, has influentially argued that the most that we could expect a transcendental argument to show is that a given phenomenon is inconceivable without the *belief* that condition x obtains, rather than its being inconceivable without condition x actually obtaining<sup>11</sup>.

Whilst this objection, if sustainable, is a major difficulty for many transcendental arguments, it should not overly worry any proponent of the arguments which are to be discussed in the following sections and chapters. To see this it is useful to distinguish between different kinds of transcendental argument<sup>12</sup>. A distinction can be made between *truth-directed*, *belief-directed*, and *experience-directed* transcendental arguments. These distinctions can be fleshed out with reference to the Body Claim. A truth-directed transcendental argument would claim

---

<sup>11</sup> Also see Körner, 1969. One philosopher who disputes Stroud's objection is Peacocke, 1989.

<sup>12</sup> Here I follow Stern, 2000, pp.10-11. His own categorization, in turn, borrows from Peacocke, 1989, pp.4-5, and Cassam, 1997, p.33.

that it is a transcendental condition of self-consciousness that one be embodied. A belief-directed transcendental argument would argue that it is a transcendental condition of self-consciousness that one believe oneself to be embodied. Finally, an experience-directed directed transcendental argument would argue that it is a transcendental condition of self-consciousness that one experience oneself as embodied. The present point is that it is only to truth-directed transcendental arguments that Stroud's criticisms apply. Belief-directed and experience-directed transcendental arguments are not affected. Since the arguments discussed in the following sections are either experience or belief-directed transcendental arguments, Stroud's well known criticism is irrelevant to their success or failure<sup>13</sup>.

We can now return to the identity argument, which is an experience-directed argument. Cassam begins by pointing out that, "a self-conscious subject must be capable of thinking of her experiences as *her* experiences, that is, of self-ascribing them" (pp.118-9). This is, he claims, "at least one important element of what might be described as our intuitive notion of self-consciousness" (p.119). Of course, this self-ascription cannot be merely accidental, as in Perry, 1979, but must be '*qua* subject'. Cassam then claims that it, "is a quite general truth that the ascription of experiences to an identical subject turns on the existence of some means of distinguishing the subject of such ascriptions as one object among others, because it is an even more general truth that a thinker will not count as having latched on to a particular item in the world and predicated something of it unless she has a capacity to distinguish the object of her judgement from all other things" (pp.122-3). One way of distinguishing oneself from all other things involves knowing "what *sort* of thing one has referred to" (p.123). But this kind of knowledge does not appear to be a transcendental condition on self-consciousness, since we can consider, "the Cartesian dualist who regards the persisting subject of her thoughts as an immaterial substance. This belief may well be philosophically indefensible...but this surely has no bearing on her ability to think first-personally" (p.127). That is, both dualists and materialists are self-conscious, even though at least one of them does not appear to know what sort of thing she is. So there must be a way of distinguishing oneself from all other things that does not presuppose that one know what sort of thing one is. The answer is that,

---

<sup>13</sup> In fact, this is only partly true. It will emerge that due to the peculiarities of the notion of self-awareness, if Cassam's identity argument is successful it implies the success of the corresponding truth-directed argument.

“awareness of the object [oneself] as a shaped, located, and solid ‘articulated unity’ is what puts S in a position to ‘isolate’ it and predicate something of it” (pp.136-7). So, “The dualist satisfies a substantive ‘knowing which’ requirement on self-reference because and only because she is intuitively aware of that to which she ascribes her experiences as an articulated physical unity” (p.140). Thus, “Awareness of the subject of one’s experiences as something with a determinate shape, as well as solidity and location, is a transcendental condition of consciousness of self-identity because it is in being aware of one’s spatial properties that one satisfies the discrimination requirement on self-reference” (p.142). This argument is complex, but can be, I think faithfully, represented in the following way:

- 1) Self-consciousness involves the self-ascription of mental predicates, *qua* subject.
- 2) The self-ascription of mental predicates *qua* subject involves having thoughts about oneself *qua* subject.
- 3) It is not possible to think about an object unless one knows what thing one’s thought is about.
- 4) Knowing what object one’s thought is about is a matter of being able to distinguish it from all other things.
- 5) It is possible to have thoughts about oneself *qua* subject even if one does not know (or truly believe) what sort of thing one is.
- 6) The only way in which one can distinguish an object from all other things when one does not know what sort of thing it is, is through an experiential awareness of it as shaped, located and solid.
- 7) Being experientially aware of something as shaped, located and solid is being experientially aware of it as a physical object.  
So,
- 8) Being experientially aware of oneself *qua* subject as a physical object is a transcendental condition of self-consciousness.

Each of these steps could use a little explanation. (1) is a consequence of the fact that self-consciousness is being understood as the capacity to think self-consciously. Thinking self-consciously is the capacity to think ‘I’-thoughts, and this means being able to ascribe mental predicates to oneself, as Cassam says, *qua* subject. (2) simply

makes the point that a subject who can self-ascribe predicates in thought is able to entertain thoughts about themselves. ‘I’-thoughts are thoughts about oneself. (3) expresses the seeming truism that there is a sense in which one must know what it is that one is thinking about. If one does not know what one is thinking about then one cannot be thinking about *it*. (4) gives a particular gloss on what it means to say that one knows what it is that one is thinking about. (5) appears to follow from the fact that both the dualist and the materialist, at least one of whom does not appear to know what kind of thing he or she is, can think self-consciously. (6) explains how it is that one is able to do this – via perceptual acquaintance. (7) gives a plausible sufficient condition of being aware of something as a physical object. (8) can clearly be read as the claim that bodily self-awareness is a transcendental condition of self-consciousness. Cassam’s argument is, of course, deeply influenced by a line of thought found in Strawson,

It is a quite general truth that the ascription of different states or determinations to an identical subject turns on the existence of some means of distinguishing or identifying the subject of such ascriptions as one object among others. Applying this general truth to the case before us, we may say, in Kant’s terminology, that the possibility of ascribing experiences to a subject of experiences and hence the possibility of self-ascription of experiences requires that there be some “determinate intuition” corresponding to the concept of a subject of experiences (Strawson, 1966a, p.102).

Despite its heritage, there are serious problems with Cassam’s identity argument. The first is that, taken together, (3) and (4) amount to an acceptance of Russell’s Principle, about which I have already expressed serious doubts. Since I have already argued against Russell’s Principle, I shall not labour the point here. It is, however, important to note that Cassam’s position is peculiarly unstable in this respect. For, if Russell’s Principle is correct, there is some reason for thinking that (2) will not follow from (1). For an acceptance of Russell’s Principle can lead us, as we saw in the last chapter, to deny that self-consciousness requires the capacity for self-reference<sup>14</sup>. Thus, in Cassam’s identity argument we see the ambiguous role of Russell’s Principle in full effect. On the one hand, the argument explicitly relies on Russell’s Principle, on the

---

<sup>14</sup> One might respond to this worry by claiming that Russell’s Principle can be made consistent with a tight conceptual connection between self-consciousness and self-reference if one thinks that a subject can identify descriptively itself as ‘the thinker of this thought’ (see chapter three, n.22). But this move would deprive Cassam’s argument of premise (6).

other the proponent of the argument must feel some pressure to reject Russell's Principle.

The second problem concerns the relationship between Cassam's argument and Cartesian dualism. Cassam's project is approached with a sharp distinction between metaphysics and epistemology. His view is that bodily-self-awareness is a transcendental condition of self-consciousness, and that this is a purely epistemological/conceptual claim. It neither rests on, nor entails, any substantive metaphysical doctrine. That is, the claim is experience-directed not truth-directed. However, it may well figure as one consideration in an argument against a dualist conception of the self. This is expressed in the following,

The issue is not whether the subject is a physical thing but whether it is experienced as a physical thing. The claim that the bodily self is the presented subject of experience would be compatible with maintaining that the subject is in fact non-physical (Cassam, 1997, p.58)

The problem with this is that Cassam's argument tacitly relies on the claim that the self is a physical object. Being aware of oneself as shaped, located and solid could only be a way of distinguishing oneself from all other things if one is, in fact, shaped, located and solid. If, say, one is a non-physical soul, then being aware of something as shaped, located and solid, could not be a way of distinguishing oneself from all other things. But this means that, pace Cassam, the argument will fail if Cartesian dualism is true. Cassam raises what is essentially this objection himself but his reply is inadequate as it *assumes* that bodily-awareness is bodily-self-awareness. He writes,

The point...is that the bodily self is not a 'mere' object but a 'subject-object'; it is the presented subject of perception and thought, so awareness of its boundaries is at the same time awareness of one's own boundaries *qua* subject of perception of thought. (Cassam, 1997, p.142)

That is, there are independently plausible reasons to think that bodily-awareness is bodily self-awareness, so awareness of the body is awareness of the presented subject. But this is, dialectically speaking, a dubious move. Cassam's argument is intended to show that bodily self-awareness is a transcendental condition of self-consciousness. Since this is the case, he is not at liberty to assume that the self-conscious subjects he takes his claim to be true of are such as to enjoy bodily self-awareness. For this is exactly what is to be shown.

Even if such a move were acceptable, it is quite unclear how it is relevant. For, supposing that the body is ‘the presented subject of perception and thought’, the question remains as to whether things are as they are presented. If they are, then the self is a physical object. If they are not, then it is obscure how being aware of something which is presented as the self, but which is not in fact the self, could suffice as a way of distinguishing the self from all other things. For one is precisely *not* aware of the self. So, again, it seems that the argument will only convince those who believe that the self is a physical object. Not only that, since this is a transcendental argument attempting to spell out the transcendental conditions of self-consciousness *per se*, it will only convince those who believe that it is necessarily so that the self is a physical object.

Another way of putting this point, is that whilst Cassam’s argument is experience-directed, the peculiarities of the notion of self-awareness are such that this experience-directed argument could only be successful if the corresponding truth-directed argument were successful also. That is, experience of one’s body as oneself could only be a transcendental condition of self-consciousness if one’s body being oneself were such a condition. And it is this claim that has not been argued for.

All in all it must be concluded that Cassam’s Identity argument is a failure. The major flaw in the argument is its unquestioning reliance on Russell’s Principle. Even if Russell’s Principle is accepted, there is reason to think that the identity argument is not as straightforward as Cassam believes. In the following section I give another reason for thinking that Cassam’s argument should be rejected. There I argue that bodily-awareness is *not* a form of self-awareness. Since we are self-conscious beings, and we do not enjoy anything that might be described as bodily self-awareness, any transcendental argument which concludes that bodily self-awareness is a transcendental condition of self-consciousness is bound to fail.

#### **4.4 Bodily-Awareness and Self-Awareness<sup>15</sup>**

The idea under consideration is that bodily-awareness is a form of self-awareness. This bodily-awareness as self-awareness claim can be evaluated independently of any transcendental arguments. What is required is a reasonably plausible set of conditions

---

<sup>15</sup> The arguments given in this and the following section owe a great deal to Martin, 1995 and 1997.

that a form of awareness must meet in order to count as a form of self-awareness. We can then go on to ask whether bodily-awareness is such as to satisfy these conditions. The best attempt to give such a set of conditions is made by Cassam,

For a form of awareness to count as awareness of something *qua* subject is for it to be (i) awareness of it as one's point of view on the world, (ii) awareness of it as a bearer of sensations, and (iii) awareness on the basis of which it is possible to make first-person statements that are immune to error through misidentification relative to 'I' (Cassam, 1997, p.77)

For the sake of argument, I grant that these three conditions are clearly specified, that they are necessary conditions of self-awareness, and that they are met by bodily-awareness. My claim is that these conditions are not jointly sufficient conditions of self-awareness. There is another necessary condition of self-awareness which is clearly not met by bodily-awareness. It follows from this that bodily-awareness does not qualify as a form of self-awareness<sup>16</sup>.

I take, as my starting point, the claim that at least the central cases of self-consciousness involve self-reference. It is a priori that at least those central cases of 'I'-thoughts refer to whoever produces them. Furthermore, this is no accidental feature of 'I'-thoughts, rather it is part of the concept of the first-person. When I think about myself in the first-person it is both a priori and manifest to me that I am thinking about myself. Now, if there is such a thing as self-awareness it is compelling to think that it serves to ground at least some central first-person judgements. After all it is difficult to see what other reason one could have for postulating such a thing. But this means that self-awareness must ground judgements which have a priori guaranteed self-reference. It must ground judgements which are manifestly about myself. But if a form of awareness is to ground judgements which are manifestly and a priori about object O, then that form of awareness must manifestly and a priori be an awareness of O. Applying this in the present case, if bodily-awareness is to count as a form of self-awareness not only must it be true that the object of bodily-awareness (which we can assume is the body<sup>17</sup>) is the self, but this must be a priori and manifest to the subject of such an awareness. Self-awareness is such that it should not be

---

<sup>16</sup> Cassam disagrees, writing "While the possibility that there are independently plausible further conditions cannot be ruled out, it would be unwise...to expect too much from this line of thought." (Cassam, 1997, p.77).

<sup>17</sup> But see Martin, 1995, and, 1997, for a note of caution on this point.

rationally open for the subject to wonder whether the object he or she is aware of is him or herself.

The first thing to note here is that the bodily-awareness as self-awareness claim, pace Cassam, relies on an a priori identification of the body and the self. An object could only satisfy all the conditions on being presented as the self if it is, in fact, the self<sup>18</sup>. Thus, the claim rests on the denial of Cartesian dualism. It is no surprise then, that Cassam's argument for the version of the Body Claim which involves bodily self-awareness tacitly relies on the denial of dualism. Since Cartesian dualism isn't exactly a popular theory these days, this might not be thought to be a serious problem. But there is another, much more pressing, difficulty.

If bodily-awareness is to count as self-awareness, it must be manifest to the subject of such an awareness that the object so observed is the self. But this condition is not met in the case of bodily-awareness. It is perfectly reasonable to wonder whether the object with which I am presented through bodily-awareness is the self, or something closely related but non-identical to the self. This is supported by the following observation. Self-ascriptions based on bodily-awareness can be expressed linguistically in two ways. One can say 'I have crossed legs' or 'I ache all over', but one can equally well say 'My legs are crossed' or 'My body aches all over'. This substitution is *always* possible<sup>19</sup>. The point is that it looks plausible to say that the content of the perceptions one has of one's own body in bodily-awareness are of the form 'My body is *F*', rather than the form 'I am *F*'. That is, through bodily-awareness my body *is* presented to me as my body, but not as myself.

A plausible account of what self-awareness must be tells us that when I am aware of myself, I know it is myself that I am aware of in this way. But this just isn't true of bodily-awareness. When I perceive my body, through bodily-awareness I *do* know that it is my body that I perceive. But it is clearly open for me to doubt that my body is myself. Bodily-awareness does not ground judgements which are manifestly and a priori about the self. Therefore, bodily-awareness does not constitute a form of self-awareness.

---

<sup>18</sup> It is a virtue of Bermúdez, 1998, p.135, and de Gaynesford, 2002, that they rest their arguments for the bodily-awareness as self-awareness claim on the premise that 'the self is embodied'. As will become clear in the following discussion, this is not enough to save bodily self-awareness.

<sup>19</sup> Note that one can't do this for mental self-ascriptions or non-mental self-ascriptions that are not based on bodily-awareness. 'My brain thinks it is going to rain' is as nonsensical as is 'My body was born in 1976'.

It will be useful to return briefly to immunity to error through misidentification. Both Cassam and Bermudez claim that self-ascriptions based on bodily-awareness are immune to error through misidentification. Furthermore they think that this fact is strong evidence for their view that bodily-awareness is a form of self-awareness. Whilst they do not claim that it is a sufficient condition of the doctrine of bodily self-awareness, both do tend to suppose that it is the most important claim to make out in a defence of that view. For example, Bermúdez says that,

the fact that somatic proprioception [bodily-awareness] is itself immune to error through misidentification relative to the first-person pronoun provides vital support for the claim that somatic proprioception counts as a genuine form of self-consciousness [self-awareness] (Bermúdez, 1998, p.144)

The point of the present considerations is to show that this is not the case. I argued in chapter two that a judgement will be IEM if it is based on a single-object faculty. It may well be that bodily-awareness is a single-object faculty. But this is perfectly compatible with the claim that the body and the self are distinct. This point is particularly clear when we notice that even the Cartesian dualist can admit that self-ascriptions based on bodily-awareness are immune to error through misidentification. For the dualist can happily allow that bodily-awareness is necessarily an awareness of one's own body only. That does not contradict Cartesian dualism. The dualist will simply go on to claim that the single object in question is not identical to the self. But, of course, the dualist will strenuously deny that bodily-awareness is a form of introspective self-awareness.

The temptation to argue from the fact that self-ascriptions based on bodily-awareness are immune to error through misidentification is simply a new application of a temptation noted by both Wittgenstein, 1958, and Strawson, 1966a, and dubbed 'the Cartesian Illusion'<sup>20</sup>. In *The Bounds of Sense* Strawson wrote,

And now we come to the fact that lies at the root of the Cartesian illusion. It may be put as follows. When a man (a subject of experience) ascribes a current or directly remembered state of consciousness to himself, no use whatever of any criteria of personal identity is required to justify the use of the pronoun 'I' to refer to the subject of that experience. It would make no sense to think or say: *This* inner

---

<sup>20</sup> See chapter two, n.9.

experience is occurring, but is it occurring to *me*? (This feeling is anger; but is it I who am feeling it?)...It is easy to become intensely aware of the immediate character, of the purely inner basis, of such self-ascription...Thus there arises a certain illusion: the illusion of a purely inner and yet subject-referring use for 'I'. (Strawson, 1966a, pp.164-6).

Here Strawson suggests that it is a misunderstanding of the IEM of mental self-ascriptions which tempts philosophers towards the view that we are introspectively aware of the soul, i.e. the claim that there is such a thing as mental self-awareness. But there is a more economical explanation available of the IEM of introspection-based self-ascriptions. This is the, extremely plausible, claim that introspection is a single object faculty. In a similar vein the fact that self-ascriptions based on bodily-awareness are IEM does not show that bodily-awareness is a form of self-awareness. All we need to allow is that bodily-awareness is a single-object faculty; that it is necessarily an awareness of only one object, one's own body. Thus the Cartesian illusion diagnosed by Strawson has a parallel Anti-Cartesian Illusion. Consideration of the notions of self-awareness, bodily-awareness and immunity to error through misidentification tells us to avoid both in equal measure.

#### **4.5 Bodily Sensations**

There is another argument for the bodily-awareness as bodily-self-awareness claim which, if sound, should lead us to re-evaluate the considerations of the previous section. This argument concerns the nature of bodily sensations, and is due to Brewer, 1995. In a way, Brewer's argument is more ambitious than Cassam's, in that it attempts to derive the metaphysical conclusion that the body is the self. His argument is summarized in the following,

In bodily-awareness, one is aware of determinately spatially located properties of the body that are also necessarily properties of the basic subject of that very awareness...a psychological property of oneself is physically located in or on the body, as a property of the body. Therefore...the animal body *is* the conscious mental subject of bodily awareness (Brewer, 1995, p.300)

The idea in play here is that bodily sensations are intrinsically spatial in character since they are perceived as located. They are perceived as located properties of the body. This idea is familiar, and I offered a brief defense of it above. Furthermore,

claims Brewer, bodily sensations are mental properties. They are necessarily properties of the subject of experience. It follows that the body is the subject of experience, and in bodily-awareness one is aware of the body as the subject of experience.

The argument is strikingly simple, but what are we to make of it? I have already defended the first claim, and so do not intend to challenge it on that account. The question to answer is whether bodily sensations (conceived of as located, subjective and mind-dependent properties) are necessarily properties of the subject of experience. The problem for Brewer is that we have yet to be given a reason to answer this question in the affirmative. As Martin, 1997, p.135, has pointed out, there is an analogy to be drawn here with colours. On certain views of colours, to perceive an object as red is to perceive redness in that object. Furthermore, an object is red if and only if it is being perceived as red. That is, according to some philosophers, we attribute subjective, mind-dependent, properties to objects distinct from ourselves. So, the fact that we attribute subjective properties to our bodies does not show that the body is, or is experienced as, the self. The claim that Brewer has not made out is that it is impossible for the body to be the possessor of subjective properties without it thereby being the self. As Gardner writes, in a slightly different connection,

Although sensation may be taken to show the existence of a point at which the inner touches the outer and, perhaps, the existence of irreducibly psycho-physical phenomena, it does not supply the idea of mind as incarnate in and, as Descartes put it, coextensive with body (Gardner, 1994, p44)

#### **4.6 Experience and Belief**

I have been arguing against the claim that bodily self-awareness is a transcendental condition of self-consciousness. Not only have I been arguing that experiencing oneself as a body is not a transcendental condition of self-consciousness, I have maintained that experience of oneself as a body is not something that we actually enjoy. All this might be accepted and yet it be thought that something akin to the strong (bodily self-awareness) version of the Body Claim is correct. That is, it might be claimed that it is a transcendental condition of self-consciousness that one *believes* that one is a physical object. Whilst one is not required to have experience of oneself as a body, one must believe oneself to be so. The quick response to this belief-

directed argument is to point out that there are, and have been, actual people who are self-conscious in the relevant sense, but who do not believe, or have not believed, themselves to be physical objects. We should not want to prove a priori that the Cartesian Meditator lacks the resources to think his famous ‘I’-thoughts<sup>21</sup>.

But maybe a more subtle approach can be taken. Campbell suggests that his own position entails that to be self-conscious you must “think of yourself as a physical thing” (Campbell, 1997, p.660). If this means believing yourself to be a physical thing I think it is false for the reason given above. But I think it doubtful that this is what Campbell does mean. In Campbell’s view, ‘thinking of oneself as a physical object’ is a matter of having a conception of oneself as a physical object. And this may well be compatible with failing to believe that one is a physical object. Especially given that such a conception may be very much implicit<sup>22</sup>. Campbell’s view is that self-consciousness requires that one grasp,

one’s own causal structure...[this involves] grasp of the idea that one’s later states causally depend, in part, on one’s earlier states...[and] the idea that one can function as a common cause of various correlated events around one (Campbell, 1994a, pp.1-2)

This kind of causal structure, according to Campbell, is “characteristic of physical objects” (Campbell, 1993, pp.92-3). More than this, to be thinking of something as a physical object, one must think of it as having this kind of causal structure. But it should be noted that even if we accept that to think of something as a physical object one must think of it as causally structured, it is not obvious that thinking of something as causally structured means that one must be thinking of it as a physical object. Thus, to show that self-consciousness requires one to think of oneself as causally structured, is not yet to show that self-consciousness requires one to think of oneself as a physical object. It is not obviously incoherent to claim that thinking of oneself as causally structured could be a matter of thinking of oneself as being mentally causally structured, and also as ‘having’ a body through which one can function as a common cause of events around one. Thinking of oneself as causally structured in Campbell’s sense would appear to be perfectly compatible with thinking of oneself as a Cartesian soul. Thinking of oneself as causally structured may be thinking of oneself as a thing,

---

<sup>21</sup> See Cassam, 1997, p.128.

<sup>22</sup> See Campbell, 1997, p.662.

but it need not be thinking of oneself as a physical thing. Thus, I don't think that Campbell's view should be regarded as an attempt to put forward a belief version of the Body Claim<sup>23</sup>.

#### 4.7 Interlude

We have arrived at a point at which it is convenient to stop and take stock of the situation. The fact that there is a tight conceptual connection between self-consciousness and self-reference implies that a self-conscious being must satisfy whatever conditions there are on self-reference. Cassam has argued that a transcendental condition of self-reference is that one is experientially aware of oneself qua subject as a physical object. If this were true, it would show that the third, strongest, form of apparent presence in the world is a transcendental condition of self-consciousness. However, Cassam's argument fails. Experience of oneself qua subject as a physical object is not even something that we enjoy, and so it cannot be a transcendental condition of self-consciousness. Furthermore, there is reason to think that it is not a transcendental condition of self-consciousness that one believe oneself to be, or think of oneself as, a physical object.

But all is not lost for the proponent of the Body Claim. Even if the third grade of apparent presence in the world cannot be established, maybe the second grade can. That is, maybe it is a transcendental condition of self-consciousness that one experiences (or believes) oneself to 'have' a body. The remainder of the thesis will be concerned with arguments that attempt to show that everyday bodily-awareness is a transcendental condition of self-consciousness. I begin with considerations that can be drawn from the work of Strawson. I argue that although Strawson's own arguments are not successful, they point us in a fruitful direction. I go on to formulate 'the self-location argument', which does manage to find an acceptable transcendental condition on self-consciousness, albeit a rather minimal one. However, the self-location argument does not get us so far as the Body Claim. Before we can endorse the Body Claim we need to supplement the self-location argument. I consider what I take to be

---

<sup>23</sup> Campbell often appears to offer transcendental arguments. However I think that this appearance is somewhat misleading. His work seems to me to be better thought of as mapping out relations between concepts *given certain empirical assumptions*. For instance he writes, "Those who have experiences of which we know nothing may be able to think in ways of which we know nothing", (Campbell, 2002, p.1).

the two most plausible ways of doing this. These are ‘the solidity argument’ and ‘the action argument’. However, it will be shown that if the self-location argument is extended into either the solidity argument or the action argument, it fails.

Something of positive value can be salvaged out of these negative results, however. For, unlike the third grade, the second grade of apparent presence in the world is one that we ourselves actually enjoy. And both the solidity and the action arguments, if supplemented with certain empirically grounded truths about our own situation, have some merit. This is, I suggest, what accounts for their initial attraction. Of course, taken in this way, the arguments no longer count as properly transcendental. They no longer map out strictly transcendental connections between concepts or conceptual capacities. But, it will be argued, they are still of interest, both in the journey they take us on, and in the more modest conclusions they reach.

## 5. Strawson

In the previous chapter I argued against the strong version of the Body Claim, which states that bodily self-awareness is a transcendental condition of self-consciousness. The weak version of the Body Claim states that bodily-awareness is a transcendental condition of self-consciousness. The aim of the present chapter is to unearth and critically evaluate those arguments for this view that can be found in the work of Strawson. This is appropriate, since it was the publication of Strawson's *Individuals* and *The Bounds of Sense* which did much to popularise the use of transcendental arguments, particularly those which aim to state transcendental conditions of self-consciousness. There are at least three arguments to be found in Strawson's work which might be used by the proponent of the Body Claim. A related argument is put forward by Bennett, 1966. I shall be arguing that none of these are convincing. However, as we shall see in the next chapter, one of Strawson's arguments contains some materials which can be reused in a better argument: the self-location argument.

### 5.1 The Objectivity Condition

Strawson, 1966a, Pt.2, Ch.2, offers at least two arguments for the claim that a self-conscious subject must satisfy what he calls 'the objectivity condition'. Strawson, 1959, Ch.2, argues that a subject can only satisfy the objectivity condition on the supposition that that subject enjoys spatial experience. If these claims can be sustained we are, it might be thought, well on our way to establishing the Body Claim. For there are arguments (specifically, the action argument, and the solidity argument, to be discussed in chapters seven and eight respectively) which purport to show that any self-conscious subject that enjoys spatial experience is such as to enjoy embodied experience. We must, then, enquire into this argument. In fact, we shall only here be concerned with the first step. For, as I shall be arguing, Strawson fails to make a convincing case for the claim that satisfaction of the objectivity condition is a transcendental condition of self-consciousness.

The first task is to clarify the objectivity condition itself. Strawson actually defines the objectivity condition in two non-equivalent ways. The first of these is in terms of a 'non-solipsistic consciousness'. A subject satisfies the objectivity condition

if it has a non-solipsistic consciousness. Strawson describes the relevant notion as follows,

I shall mean by a non-solipsistic consciousness, the consciousness of a being who has a use for the distinction between himself and his states on the one hand, and something not himself or a state of himself, of which he has experience, on the other (Strawson, 1959, p.69)

So, on this reading of the objectivity condition (OC1), the claim is that self-consciousness requires that one have a use for the distinction between oneself and one's states, and that which is not oneself or one of one's states<sup>1</sup>. Two things should be noted about this reading. The first is that the phrase 'must have a use for the distinction' is rather vague, and could be cashed out in a variety of ways. On a strong reading, it might mean 'must actually (and correctly) apply the distinction'. Alternatively it could mean 'must believe there to be such a distinction'. Third, it might mean 'must have experience such as to justify the application of such a distinction'. Finally, it could mean 'must have the conceptual resources required to understand/make the distinction'. I will not, here, adjudicate between these readings, but simply suggest that something like a combination of the third and fourth readings is what Strawson has in mind. The idea is that if one is self-conscious one must have the conceptual sophistication to make judgements which involve the self/world distinction. Furthermore, one could only have this capacity if one's experience were such as to justify judgements that involve that distinction. Strawson's view seems to be that one's conceptual repertoire is constrained by one's experiential situation. The range of one's conceptual repertoire cannot vastly outstrip one's experience. We can, for the sake of argument, accept this as a premise<sup>2</sup>.

The second point to bear in mind is that OC1 seems to be concerned with mental self-ascription rather than both mental and physical. The states to which Strawson is referring are experiential states, belief states etc. The idea is that a self-conscious subject must be able to distinguish between its experiences and what those experiences are experiences of (or between its beliefs and what those beliefs are beliefs about). For example, the capacity to entertain the self-conscious thought 'I can

---

<sup>1</sup> This also appears to be the understanding of the objectivity condition in play in Bennett, 1966, §§32, 33 & 51.

<sup>2</sup> This hierarchy of levels at which OC1 can be pitched is also discernable with regards to OC2. I will refrain from making the same point again.

see an ostrich' presupposes the capacity to distinguish between one's experience of an ostrich and the actual ostrich.

Strawson's second way of understanding the objectivity condition (OC2) is as follows: a subject who satisfies the objectivity condition is one whose experience is more than just a 'pure sense-datum experience'. On this reading Strawson's claim is that a self-conscious subject cannot have a purely sense-datum experience. A purely sense-datum experience is characterised as one in which there is,

no distinction to be drawn between the order and arrangement of the objects (and of their particular features and characteristics) and the order and arrangement of the subject's awareness of them (Strawson, 1966a, p.99)

A purely sense-datum experience is entirely subjective in character. We might think of it as consisting in nothing but a series of auditory and olfactory experiences<sup>3</sup>. The important point is that a purely sense-datum experience is one which does not allow for conceptualisation in terms of the seems/is distinction. On this reading, Strawson's claim is that a self-conscious subject could not have experience of just this type. Another way of putting OC2 is that a subject who satisfies the objectivity condition must have a use for the notion of existence unperceived. This is a construal of the objectivity condition which has been picked up by both Evans, 1980, and Cassam, 1997. For example, Cassam writes that for one to satisfy the objectivity condition,

is for one to be in a position to think of it [one's experience] as including perceptions of objects in the 'weighty' sense, that is, particular items which are capable of being perceived and of existing unperceived (Cassam, 1997, p.28)

The notion of existence unperceived comes in a package with the seems/is distinction. If one is in a position to think of the objects of one's experience as having an order and arrangement different to the order and arrangement of one's experiences, then one must be in a position to think of the objects of one's experience as enjoying some sort

---

<sup>3</sup> Strawson includes "red, round patches, brown oblongs, flashes, whistles, tickling sensations, smells." (Strawson, 1966a, p.99). This list is unsatisfactory for two reasons. First, as I argued in chapter four, bodily sensations are felt as located and are not entirely subjective states. Second, since the aim is to argue for the thesis that a self-conscious subject must have spatial experience, it begs the question to assume that the subject has (spatial) visual experience. One should not start one's journey from the place at which one wishes to arrive!

of unperceived existence. For instance, if one experiences the temporally ordered series of smells {nice—nasty—nice}, and one has the capacity to conceive that the objects of experience might actually have been ordered {nice—nice—nasty}, then one must be in a position to think of {nice} as having existed unperceived. The converse also holds, since if something exists unperceived, its order and arrangement will necessarily differ from the order and arrangement of one's experience of it.

Strawson and his commentators tend to treat OC1 and OC2 as equivalent, but they are wrong to do so. The equivalence of OC1 and OC2 is something which is far from obvious, and would therefore need to be argued for. To see that OC2 does not entail OC1 one can suppose that one has experience solely of oneself and one's states, and that one can conceive of one's own states existing unperceived (e.g. a not-presently-conscious belief), or even of oneself existing unperceived (e.g. when one is asleep). Here, although one satisfies OC2, one does not satisfy OC1, i.e. one can distinguish between the order and arrangement of one's experience of oneself and the order and arrangement of oneself. But since one's experience is solely of oneself and one's states, one has no use for the distinction between oneself and one's states and that which is not oneself or one's states. Thus, one can have a use for the seems/is distinction without having a use for the self/world distinction. Nor does OC1 entail OC2. One might have a use for the self/world distinction, and yet be in no position to conceptualise the possibility that that which is other than oneself could exist whilst unperceived. One's experience might be so meagre as to fail to 'make room for' the distinction between seeming and being. Thinking of something's existence as dependant upon one's perception of it is not the same as thinking of it as a state of oneself (consider, for example, the case of colour). So, having a use for the self/world distinction does not entail having a use for the seems/is distinction.

Of course if Strawson is correct in claiming that a non-solipsistic consciousness (or a purely sense-datum experience) is not possible for a self-conscious subject, we will have to revise these claims. For such a subject could not have the sorts of experiences that I have just assumed are possible. Any experience would have to make room for the self/world (or the seems/is) distinction. But this is Strawson's conclusion and cannot be assumed at the outset. Since these different versions of the objectivity condition are in play, we must be careful in evaluating any argument for the claim that self-consciousness requires satisfaction of the objectivity

condition. We must be clear as to exactly which version of the objectivity condition, if any, is established.

## 5.2 Sketch of an Over Hasty Argument

It might be thought that there is a simple way of establishing that self-consciousness requires satisfaction of the objectivity condition (OC1). Self-consciousness involves the thinking of 'I'-thoughts. As such, self-conscious thoughts involve the concept *I*. That is, any self-conscious being can be attributed the concept *myself*. But a concept serves as a way of dividing up the world into two: that which falls under the concept and that which does not. Thus, if I possess the concept *I*, then I ipso facto possess the (complex) concept *not-I*. But if one is in possession of these two concepts then one satisfies the objectivity condition (OC1). If one is in possession of the concepts *I* and *not-I* one has a use for the distinction between oneself and one's experiences, and that which is not oneself or one's experiences, of which one has experience (i.e. one is a subject with a non-solipsistic consciousness). Thus, a self-conscious subject must satisfy the objectivity condition.

But the last step of this argument is too quick. Satisfaction of the objectivity condition (OC1) does not follow from a subject's possession of the concepts *I* and *not-I*. A non-solipsistic consciousness is defined as one which has a use for the distinction between itself *and its experiences* and that which is not itself *or its experiences*. All we have secured is that a self-conscious being must have a use for the distinction between itself and that which is not itself. But a subject's experiences are not (are not identical to) itself. Thus, the use a subject has for the distinction between *I* and *not-I* may be exhausted by its capacity to distinguish between itself and its experiences. That is, having the capacity to make judgements using the concept *not-I* is not yet to have the capacity to make judgements involving that which is neither oneself nor any of one's experiences.

## 5.3 Strawson's Objectivity Argument

In presenting his objectivity argument Strawson claims to be offering a reconstruction of Kant's transcendental deduction and the related refutation of idealism. Whether Kant can really be interpreted in this way is doubtful, but the objectivity argument

is very much interesting enough in its own right to be studied independently of its supposed Kantian heritage<sup>4</sup>.

Strawson's objectivity argument attempts to show that any self-conscious subject must satisfy the objectivity condition. Strawson begins by making the connection, with which we are familiar, between self-consciousness and self-ascription. This he does via the notion of the unity of consciousness, "The condition under which diverse representations may be said to be united in a single consciousness is precisely the condition, whatever that may be, under which a subject of experiences may ascribe different experiences to himself, conscious of the identity of that to which these different experiences, at different times, belong." (Strawson, 1966a, pp.95-6). Now, in order to self-ascribe experiences, one must be able to think of experiences *as* experiences. Strawson makes this claim by asserting that experience is necessarily self-reflexive, "What is meant by the necessary self-reflexiveness of a possible experience in general could be otherwise expressed by saying that experience must be such as to provide room for the thought of experience itself." (Strawson, 1966a, p.107). Finally, Strawson claims that possessing this capacity to think about experience presupposes the satisfaction of the objectivity condition, "For 'This is how things are (have been) experienced *by me* as being' presupposes 'This is how things are (have been) *experienced* as being'; and the latter in turn presupposes a distinction, though not (usually) and opposition, between 'This is how things are experienced as being' and 'Thus and so is how things are'" (Strawson, 1966a, p.108). Strawson links this explicitly to the objectivity condition in the following, "The point of the objectivity-condition is that it provides room for this thought. It provides room, on the one hand, for 'Thus and so is how things objectively are' and, on the other, for 'This is how things are experienced as being'; and it provides room for the second thought *because* it provides room for the first. (Strawson, 1966a, p.107). Thus, we have moved from the bare possibility of self-consciousness to the notion of a unified self-conscious subject satisfying the objectivity condition. We can reconstruct this argument as involving three central steps<sup>5</sup>:

---

<sup>4</sup> See Aquila, 1979, for doubts about the Strawsonian interpretation of Kant's transcendental deduction and, in particular, the refutation of idealism.

<sup>5</sup> It will be noticed that the following reconstruction makes no reference to the unity of consciousness. Indeed, I suggest that the unity of consciousness is a red herring in the context of the objectivity argument. This point is recognized in Rorty, 1970, p.214. For different conceptions of, and attitudes towards, the unity of consciousness, see Brentano, 1973, Nagel, 1971, Henrich, 1989, Hurley, 1994, and Dainton, 2000.

- (1) Self-consciousness involves the self-ascription of experiences.
- (2) The self-ascription of experiences is possible only because experience necessarily makes room for the thought of experience *as* experience.
- (3) Thinking of experience *as* experience requires satisfaction of the objectivity condition.

This argument for the objectivity condition relies, claims Strawson, on less than it might. That is, it does not take advantage of the full transcendental conditions of self-consciousness. Strawson makes a distinction between personal and transcendental self-consciousness. Transcendental self-consciousness is identified with the necessary reflexiveness of experience and is,

Something less than, though entailed by, the satisfaction of the full conditions of the possibility of empirical self-ascription of experiences...which yet really does constitute the essential core of personal self-consciousness. (Strawson, 1966a, p.107)

That the argument relies on transcendental self-consciousness, rather than personal self-consciousness, means that there are more resources available for a defence of the claim that self-consciousness requires the objectivity condition. We shall come back to this point in a discussion of another defence of the objectivity condition, offered by Strawson, which takes as its starting point personal rather than merely transcendental self-consciousness.

There are a host of things that can be said with regards to Strawson's objectivity argument. As it stands it is far from transparent, and there are a variety of objections that might be raised. The first thing to note, however, is that the objectivity condition that is being argued for here is OC2. The conclusion is that a self-conscious subject must be able to make the distinction between 'this is how things are' and 'this is how things are experienced as being', that is they must be in possession of the seems/is distinction. But, as we have seen, it is not obvious that possession of this distinction entails possession of the self/world distinction. This is for the reason that one's experience might be solely of oneself and one's states, and whilst one can distinguish in thought between the way things are and the way things are experienced as being, this distinction could be exhausted by one's being able to distinguish in thought between how one is and how one experiences oneself as being. Thus, the

objectivity argument as presented is an argument from transcendental self-consciousness to the objectivity condition, conceived of as OC2<sup>6</sup>.

The first objection to Strawson's objectivity argument is that the abstraction from personal to transcendental self-consciousness is an illegitimate move. This objection has been made by Cassam, 1997, Ch.3. Cassam's concern is that Strawson's own view of the conditions of reference to experiential states undermines the very notion of transcendental self-consciousness. Strawson's view on experiences is the following,

states of consciousness, or private experiences...cannot be thus identifyingly referred to except as the states or experiences *of* some identified person. States, or experiences, one might say, *owe* their identity as particulars to the identity of the person whose states or experiences they are. (Strawson, 1959, p.97)

Strawson takes it that this means that in order to know which experience one is referring to one must know which person that experience belongs to. Cassam now states his objection,

To suppose that transcendental self-consciousness can be abstracted from personal self-consciousness is...to suppose that the thought of experiences as experiences is intelligible independently of the thought that they belong to a subject who is an object among others in the world. This 'independent intelligibility thesis' is, however, precisely what Strawson's own conception of the identity of experiences appears to call into question. (Cassam, 1997, p.107).

This objection is misguided, for the following reason. Strawson's position on the identity of experiences involves, at least, two claims: (i) experiences are necessarily owned by subjects of experience, and (ii) in order to refer to an experience one must refer, or be in a position to refer, to the subject whose experience it is. The reason why the objection is misguided is that at no point does the proponent of Strawson's objectivity argument need to reject either of these two claims. The notion of transcendental self-consciousness merely represents an argumentative strategy. The

---

<sup>6</sup> There is another central aspect of Strawson's objectivity argument that I have not mentioned. For Strawson sometimes tends to argue from the possibility of experience, rather than self-consciousness. Such an argument would be far stronger than that presently reconstructed. For obvious reasons I have restricted myself to those of Strawson's claims which seem to take self-consciousness as their premise. Those aspects of Strawson's argument that concern experience simpliciter are discussed in Sacks, 2000, pp.222-228. I also ignore issues surrounding the clear anti-sceptical intent of Strawson's argument. On this see Stern, 2000, pp.137-142.

idea is that it is possible to give some of the transcendental conditions of self-consciousness without making use of all of our resources. Strawson does not imply that we can make sense of the idea of unowned experiences, or of referring to experiences without referring to their owners. Rather Strawson claims that self-conscious experience involves a certain feature, “a certain character of self-reflexiveness” (Strawson, 1966a, p.107), and that the satisfaction of the objectivity condition is a transcendental condition of *this* rather than any other feature of self-consciousness<sup>7</sup>.

Whilst step (1) of the objectivity argument is both familiar and highly plausible, steps (2) and (3) both require some explication. Concerning step (2), we can ask both what it means to say, and why we should believe, that ‘experience must be such as to provide room for the thought of experience itself’. Concerning step two we can ask why being able to think of experiences *as* experiences means that a subject must grasp the seems/is distinction.

In fact we already have the materials with which to answer the first part of the first of these questions. In the previous section I distinguished different levels at which the objectivity condition could be pitched. There I said that it was best seen as meaning that to satisfy the objectivity condition one must have the conceptual capacity to employ the self/world (or seems/is) distinction, and that one’s experience must be such as to justify the application of that distinction. I take it that we can read ‘experience must be such as to provide room for the thought of experience itself’ in the same way. That is, as being the claim that a self-conscious subject must have the concept *experience* in his or her conceptual repertoire, and that the experience of such a self-conscious subject must be such as to justify the application of that concept in judgements. This then is the second step of the argument: a self-conscious subject must have, and be justified in using, the concept of experience.

It remains to be seen what reasons there are to accept this claim. The following passage helps us to see why Strawson thinks the concept of experience indispensable,

It was agreed at the outset that experience requires both particular intuitions and general concepts. There can be no experience at all which does not involve the recognition of particular items *as* being of

---

<sup>7</sup> As Bennett rightly says, Strawson does not claim, “that self-reflexiveness could occur in the absence of personal self-consciousness” (Bennett, 1968, p.342).

such and such a general kind. It seems that it must be possible, even in the most fleeting and purely subjective of impressions, to distinguish a component of recognition, or judgement, which is not simply identical with, or wholly absorbed by, the particular item which is recognized, which forms the topic of judgement...Recognition implies the *potential* acknowledgement of the experience into which recognition necessarily enters as being one's own (Strawson, 1966a, p.100)

The idea central to this passage is, I suggest, the following. Experience is always experience *as*. Strawson construes this as the claim that experiences involves the recognition (or judgement) of itself as falling under some general concept (hence his calling it self-reflexive). Furthermore, this recognition of the experience as falling under a general concept implies the possibility of self-ascription. Recognising an experience as one's own involves making use of the concept of experience. Thus, experience must justify the application of the concept of experience.

If this is Strawson's argument, there are several concerns which must be addressed. To begin with it is left entirely obscure what is meant by experience *involving* a judgement, or having a judgement as a *component*. Second, it is not obvious that the conceptualisation of experience need involve the conceptualisation of experience *as* experience. The subject could conceptualise his or her experiences as *loud* or as *nasty*, and this would appear to satisfy the demand that they involve a conceptual component. As Harrison writes,

we would be perfectly satisfied if such an experiencer could describe his experiences, even if he were perfectly incapable of classifying them as experiences (Harrison, 1970, p.221)

If this objection is along the right lines, it is difficult to see how Strawson could substantiate his claim that experience must provide room for the thought of experience itself<sup>8</sup>. In my view this objection highlights one of the serious problems with Strawson's tendency to make claims concerning the transcendental conditions of bare experience, rather than of self-consciousness (see n.7). In several places Strawson makes the claim that experience requires the possibility of self-consciousness and thus self-ascription, for example he writes, "the possibility of experience requires the satisfaction of the basic condition of self-consciousness" (Strawson, 1966a, p.110). Commentators have been quick to point out that this would

---

<sup>8</sup> Cassam, 1997, pp.103-5, argues that this objection is unsuccessful. But his reply relies on Russell's Principle and so is not available to me.

appear to deny experience to animals<sup>9</sup>. But it must be remembered that our present concern is with the transcendental conditions of self-consciousness. We are free to assume that the only experience with which we are concerned is that enjoyed by self-conscious subjects. As such it is not clear that Harrison's objection is convincing. For self-consciousness requires the self-ascription of states. Let us take as an example, the self-ascription of auditory experiences. The subject hears a sound and thinks 'I am hearing a sound'. Now, whilst this does not explicitly employ the word 'experience', surely such a subject has a grasp of the concept of experience. Hearing is nothing if not a way of experiencing. It would seem that if the subject is a self-ascriber, then that subject must not only be able to conceptualise the sound heard as, say, loud. The subject must also be able to conceptualise the sound as a sound, that is, as an experience. Thus, we have rather a quick route to step (2) of Strawson's objectivity argument. Self-consciousness requires a grasp of the concept of experience.

We can now move on to Strawson's claim, in step (3), that being able to think of experience *as* experience requires that one satisfy the objectivity condition (OC1). In support of this claim Strawson tells us that,

"This is how things are (have been) experienced *by me* as being" presupposes "This is how things are (have been) *experienced* as being"; and the latter in turn presupposes a distinction, though not (usually) and opposition, between "This is how things are experienced as being" and "Thus and so is how things are". (Strawson, 1966a, p.108)

Is this convincing? To begin with, we can agree that, experience is necessarily *of* something. The proponent of the hypothesis of a purely sense-datum experience will argue that, grasping the concept of experience requires a grasp of the idea that experience is of something, but this does not amount to a grasp of the idea that experience is of objects which could either exist unperceived or differ in order and arrangement from experience of them. Rather, the grasp of the concept of experience that the purely sense-datum experience will involve, is a grasp of the notion of a *subjective* object. Subjective objects are things of which one can be aware but for

---

<sup>9</sup> See, for example, Harrison, 1970, p.215. Also see Sacks, 2000, p.227. The tendency to claim that consciousness required self-consciousness is, of course, already present in Kant. See Bennett, 1966, p.105-6, for reasons why we should not be overly worried about this. Rorty goes even further than Strawson, claiming that, "unless we restrict the application of the term 'experience' to language users...we will have no basis whatever for offering transcendental arguments about the possibility of experience." (Rorty, 1970, p.224).

which there is no seems/is distinction. Consider the putative subject solely aware of sounds and smells. We can agree that this subject must be able to think of its experiences, otherwise it would not be able to self-ascribe them. But we have not yet been given a compelling reason to think that the sense in which such a subject grasps the concept of an experience, as something of which it is aware, entails that it has any grasp of the seems/is distinction. The point is that thinking of something as an experience can fall short of thinking of something as a seeming, if thinking of something as a seeming involves this rich conceptual background of the seems/is distinction.

Of course Strawson would deny all of this. Approaching the same point from another direction, Strawson would point out that such an awareness of subjective objects that I have been maintaining is possible would fail to satisfy the requirement that there be a distinguishable recognitional aspect to all experience. That is, the recognition (or judgement) would be “identical with, or wholly absorbed by, the particular item which is recognised” (Strawson, 1966a, p.100). His point is that judgments of, that is self-ascriptions of, subjective objects would fail to be distinct from the subjective object itself. But it is hard to find, in Strawson’s remarks, any compelling reason to accept this. Unless we can find such a reason, we must conclude that Strawson’s objectivity argument has failed to show that a self-conscious subject must have some grasp either of the seems/is distinction or the self/world distinction.

#### **5.4 Empirically Applicable Criteria of Identity**

If the argument from transcendental self-consciousness to the objectivity condition fails, maybe Strawson’s argument from personal self-consciousness to the objectivity condition will fare better. This argument turns on Strawson’s view that experiences ‘owe their identity to the person whose experiences they are’. The idea is this: self-consciousness entails the unity of consciousness (consciousness of the possibility of self-ascribing diverse experiences), the unity of consciousness requires that different experiences could be ascribed to the self same subject, this requires that the subject has some way of distinguishing himself from other things, and this entails the objectivity condition. The vital claim is the following:

It is a quite general truth that the ascription of different states or determinations to an identical subject turns on the existence of some means of distinguishing or identifying the subject of such ascriptions as one object among others. Applying this general truth to the case before us, we may say, in Kant's terminology, that the possibility of ascribing experiences to a subject of experiences and hence the possibility of self-ascription of experiences requires that there be some "determinate intuition" corresponding to the concept of a subject of experiences; or, substituting another terminology for Kant's, we may say that this possibility requires that there be empirically applicable criteria of identity for subjects of experience. (Strawson, 1966a, p.102)

It should be clear from the preceding chapters that Strawson's 'quite general truth' is a substantive and highly controversial thesis. First, it appears to simply be an instance of Russell's Principle, second it is far from clear that there is, or even could be, such a thing as an intuition as of the subject of experiences. Even if Strawson could steer clear of these problems, however, it remains an open question as to whether the objectivity condition follows. The claim is that the self-ascription of experiences requires that the subject conceive of himself qua subject as having empirically applicable criteria of identity. Assuming that we know what it means to have 'empirically applicable' criteria of identity, it still does not seem that the subject satisfying this condition must have a use for the self/world distinction. What the criteria of identity might serve to do would be to enable the subject to distinguish himself from his experiences; himself from what are merely states of himself. If so, OC1 would not be entailed.

What about OC2? Does the claim that the self-ascription of experiences requires that the subject conceive of himself qua subject as having empirically applicable criteria of identity entail that he could understand the notion of existence unperceived? An argument for this claim might proceed as follows: a subject can only think of his experiences as belonging to a person (with empirically applicable criteria of identity) if he thinks of them as belonging to a physical object. He can only think of them as belonging to a physical object if he has the concept of a physical object, and the concept of a physical object necessarily involves the notion of existence unperceived. I don't think that this argument is very convincing. It is true that possession of the concept of a physical object necessarily involves the notion of existence unperceived; this step of the argument seems fine. The problematic step is the first, that in order to think of one's experiences as belonging to something with empirically applicable criteria of identity, one must be thinking of them as belonging

to a physical object, with ‘physical object’ understood as involving such a rich conceptual background. It is certainly not obvious that there is no room for the thought that one’s experiences are had by a subject with empirically applicable criteria of identity but which falls short of being thought of as a physical object which can exist unperceived. Such a line of thought is suggested by remarks made by Campbell, albeit in a rather different context,

there seems to be a level of thought more primitive than our thought about physical objects...This primitive level of thought is perhaps exemplified by the way in which we ordinarily think about the stars. If we are asked, as we look at the night sky and try to identify constellations, whether we think of the stars as physical objects or as points of light with no more causal significance than shadows, we may have no immediate answer. We are not really thinking of them as either (Campbell, 1994a, p.29)

The idea is that although we think of shadows or stars as having empirically applicable criteria of identity, we need not think of them as physical objects, and we need not think of them (although we in fact do) as being capable of existing unperceived. But could *all* of our experience be as of these ‘features’? There is some reason to think that if it were, we could no longer talk of empirically applicable criteria of identity. Whilst we do think of features as having empirically applicable criteria of identity (we can think of a shadow as *the same* shadow as before) it is plausible to think that the identity of features is dependent on the identity of physical objects (it is the same shadow because it is being cast by the same building)<sup>10</sup>. If this is the case, thinking of the subject of experiences as having empirical criteria of identity would appear to bring with it some notion of a physical object. Specifically, it would bring with it the notion of an object which is, in Campbell’s phrase (see chapter four) internally causally connected. That is, objects whose later states are causally dependent on their earlier states. Only with such a notion in play can criteria of identity be empirically applied. So, this objection to Strawson’s argument does not succeed.

The real problem with the argument is that we have not bridged the gap between thinking of something as internally causally connected, and thinking of it as capable of existence unperceived. We think of physical objects as possessing both of these features, but there seems no obvious reason why one could not think of an

---

<sup>10</sup> I present this argument in more detail in chapter seven.

object as being internally causally connected and yet as incapable of existing unperceived. Maybe we could think of Cartesian souls in this way. The present condition of Cartesian soul would partly depend on its condition at an earlier time, but its existence is dependent on its being perceived (by itself, or by God perhaps). The problem is that the rich conceptual background of existence unperceived is not required for the empirically adequate criteria of identity condition. All that is required is that the subject have experience as of objects that are internally causally connected. For this reason, Strawson's argument from personal-self consciousness to OC2 fails even if the concerns about Russell's Principle and self-awareness can be sidestepped. So Strawson's objectivity argument is unsuccessful: we have not yet been given a strong enough reason to accept the claim that self-consciousness requires satisfaction of the objectivity condition (however that is understood).

### **5.5 Bennett's Realism Argument**

Another argument for the claim that the self-consciousness requires satisfaction of the objectivity condition appears in Bennett, 1966, Ch.14<sup>11</sup>. Bennett's argument is related to the two of Strawson's arguments just discussed in that both claim to be inspired by Kant's transcendental deduction and refutation of idealism. Bennett's argument differs from Strawson's objectivity argument, in that it makes essential use of the notion of the unity of consciousness and that it centres not on the concept of experience or the concept of a physical object, but on the concept of the past. Briefly, the argument can be stated as follows: self-consciousness requires that one has the concept of the past, and having the concept of the past would be impossible unless one satisfied the objectivity condition.

Why should we think that self-consciousness requires that one has the concept of the past? Bennett says surprisingly little in defence of this claim. The best defence of it, however, must surely come in the form of the unity of consciousness. Bennett's reading of the unity of consciousness amounts to the constraint that it must be possible for one to be conscious of the identity of the subject of *actual* diverse experiences, so that the subject must be able to think 'I am now *F* and I was then *G*'.

---

<sup>11</sup> A discussion of Bennett's realism argument, and how it fits into the main thread of Bennett's interpretation of Kant, can be found in Strawson, 1968, pp.336-8.

Let us accept this, for the sake of argument<sup>12</sup>. There does seem to be some intuitive plausibility in the claim that self-consciousness as we know it would not be possible without a grasp of the concept of the past<sup>13</sup>. The hypothetical subject for whom the world is continually present, who has no conception of its past seems barely intelligible. The consciousness of a subject with the capacity to think ‘I exist’ but without the conceptual resources to judge ‘I existed before’, would be radically unlike the sort of temporally sensitive self-consciousness that we enjoy. As Heidegger puts it, “I *am-as-having-been*” (Heidegger, 1962, p.376).

What should we make of the second step of the argument? The second step must justify the claim that having the concept of the past would be impossible unless one satisfied the objectivity condition. The argument runs as follows. A subject can only be credited with possession of a concept if they have a use for that concept. If one did not satisfy the objectivity condition one would only have access to oneself and one’s states (Bennett’s version of the objectivity requirement is OC1). Judgements about the past can only be based on present data. So, any judgement about a past state of the subject must be based on a present recollection of that state. So, there will be a one-one mapping of judgements of the form ‘I was *F*’ with judgements of the form ‘I recollect being *F*’. But this makes the judgement ‘I was *F*’, and thus the concept of the past, redundant. So, the subject does not have the concept of the past. On the other hand, if one does satisfy the objectivity condition one has access to objective states of affairs (that which is not oneself or a state of oneself). This means that one can bring to bear several present data on a single judgement about the past. Thus, one has a complex relation between one’s concept of *I was F* and one’s concept of *I recollect being F*. So his concept of the past is not redundant.

The argument is rather compressed, but the basic idea should be clear. I want to make two points. The first is that it rests on a controversial principle of concept possession. I do not wish to argue against the view that a subject can only be credited with possession of a concept if he has a use for that concept, merely point out that it requires justification<sup>14</sup>. The second point is to question the claim that for a subject

---

<sup>12</sup> Although there is a worry about regress. If, to be counted as a self-conscious subject, one must be able to judge ‘I was *G* at *t*’ does this not mean that it must have been possible for me to judge, at *t*, ‘I am now *G*’? In other words, does it not mean that the subject must have been self-conscious at *t*? If so, it seems impossible for a subject to ever *become* self-conscious.

<sup>13</sup> Campbell, 1994b, makes a similar claim.

<sup>14</sup> G. Strawson, 1999, §VIII, argues against this principle.

who only has access to his own states, there will be a one-one mapping of judgements of the form ‘I was *F*’ with judgements of the form ‘I recollect being *F*’. Might it not be that such a subject could bring many present recollections to bear on a past judgement? He might have the thought ‘I recollect being *F*, and I recollect many instances of being *F* and then *G* in quick succession’ and on this basis form the judgement ‘I was *G*’. This seems to describe a possible situation. If it does, then there is no one-one mapping, and the subject’s concept of the past is not idle. Thus, there is no reason to deny that the subject has such a concept.

If this is correct, the realism argument fails to establish that a self-conscious subject must satisfy OC1, but might it succeed in establishing that a self-conscious subject must satisfy OC2? I don’t think so. There seems to be no more reason for saying that a subject with a purely sense-datum experience and no conception of existence unperceived must lack the concept of the past, than there is for a subject who fails to satisfy OC1. The sense-datum experiencer could bring several present experiences to bear on judgements about the past, and so could have a complex relationship between *I was F* and *I recollect being F*. So it appears that Bennett’s realism argument fails to establish either objectivity condition.

It was claimed that a self-conscious subject must be such as to satisfy the objectivity condition. However, consideration of the objectivity condition, in all of its forms, has failed to yield the promised results. Since we have not been given a satisfactory argument from self-consciousness (via the unity of consciousness) to the satisfaction of any version of the objectivity condition, there seems little reason to look at the second premise of the argument (the claim that satisfaction of the objectivity condition requires a subject to enjoy spatial experience). It will be more fruitful to move on to another argument, also due to Strawson, which might be more successful in supporting the Body Claim. For this, we must move back from *The Bounds of Sense* to Strawson’s earlier work, *Individuals*.

## **5.6 Strawson Against Dualism**

In Chapter 3 of *Individuals*, Strawson presents an argument against Cartesian dualism. This argument could be used to motivate the, weak version of, the Body Claim. For, if it can be shown that self-consciousness demands that Cartesian dualism is false, then it seems reasonable to think that it can be shown that we are necessarily embodied and

presumably must experience ourselves as such. In this section I will present and critically evaluate Strawson's argument. The argument can be summarised as follows<sup>15</sup>:

- (1) One can self-ascribe P-predicates only if one can ascribe them to other subjects<sup>16</sup>.
- (2) To ascribe a P-predicate to another subject, one must be able to identify that subject.
- (3) Subjects cannot be identified by others purely *as* subjects of experience<sup>17</sup>.
- (4) If subjects are Cartesian egos then they can only be identified *as* subjects of experience.
- (5) So, if subjects are Cartesian egos, they cannot identify other subjects.
- (6) So they cannot ascribe P-predicates to others.
- (7) So they cannot ascribe P-predicates to themselves.
- (8) Since we do ascribe P-predicates to ourselves, we are not Cartesian egos.

Since we are not Cartesian egos, it seems that we must be embodied. And for us to be able to identify other subjects and ourselves, we must experience others and ourselves as such. Thus, Strawson's argument would appear to support the Body Claim.

But this is not entirely convincing. The first thing that must be said with regards to this argument is that (1) is open to a number of different interpretations. At one stage, Strawson writes that, "One can ascribe states of consciousness to oneself only if one can ascribe them to others" (Strawson, 1959, p.100). The use of the word 'can' here suggests that what Strawson has in mind is the dependency of one sort of ability, or capacity, on another. However, one page previously, Strawson offers a rather different formulation, claiming that "it is a necessary condition of one's

---

<sup>15</sup> See Strawson, 1959, p.100. Strawson's argument is read in this way by Crowley in his doctoral dissertation, Crowley, 1980, pp.111-113. Also see Williams, 1961, p.121.

<sup>16</sup> P-predicates are 'Person-predicates'. They are to be distinguished from M-predicates, which are predicates that are, "properly applied to material bodies to which we would not dream of applying predicates implying states of consciousness." (Strawson, 1959, p.104). Ishiguro, 1980, p.66, fn.2, points out that Strawson seems to be suggesting that only persons have states of consciousness. This problem is relevant in what follows. Williams, 1961, pp.122-126, has worries concerning the coherence of the distinction between P-predicates and M-predicates.

<sup>17</sup> "Can we not identify such a subject as, for example, 'the subject that stands to that body in the same special relation as I stand in to this one'...? But this suggestion is useless. It requires me to have noted that *my* experiences stand in a special relation to body M...But what is the word "my" doing in this explanation?" (Strawson, 1959, p.101). The word 'my' cannot occur, since the aim is to give an account of the conditions under which we can self-ascribe P-predicates, i.e. it would be circular.

ascribing states of consciousness, experiences, to oneself, in the way one does, that one should also ascribe them, or be prepared to ascribe them, to others who are not oneself.” (Strawson, 1959, p.99). Here, we have no talk of capacities but rather of actual behaviour. The qualification made by Strawson (being ‘prepared to ascribe them’) seems to suggest a willingness to so ascribe, or maybe that one is epistemically justified in so ascribing. It should be clear that the ‘capacity version’ is the weaker claim. For, whilst it *might* be true that one can only have the capacity to self-ascribe or other-ascribe states of consciousness if one actually does self-ascribe or other-ascribe states of consciousness, this is not immediately obvious. It seems that one may have the capacity to do something yet simply refuse to do so. Taking him to have the capacity version in mind, how does Strawson defend this dependence claim? It is based on a perfectly general thesis about the nature of predicates. In an important footnote, Strawson writes,

The main point here is a purely logical one: the idea of a predicate is correlative with that of a *range* of distinguishable individuals of which the predicate can be significantly, though not necessarily truly, affirmed. (Strawson, 1959, p.99, fn.1)

The idea is that unless one has grasped the possibility of affirming a predicate of a range of individuals, one has not grasped that predicate. This might be read as the claim that one does not have the capacity to think that  $a$  is  $\phi$  unless one has the capacity to think that  $b$  is  $\phi$ . But it cannot be read as the claim that one cannot actually think that  $a$  is  $\phi$  unless one has actually thought that  $b$  is  $\phi$ , for then one would never be able to think that anything is  $\phi$ ! Thus it only supports the capacity version.

We have encountered this principle before, in Evans’ Generality Constraint (see chapter three). Indeed, Evans took himself to be elaborating and expanding upon the point made already by Strawson<sup>18</sup>. If we accept the Generality Constraint, is it possible to show that ‘one can ascribe states of consciousness to oneself only if one can ascribe them to others’? The question arises, ‘Other what?’. Would it constitute a grasp of the predicate ‘is in pain’ if one were able to ascribe it to both oneself and the various non-human animals one might have experience of? It seems difficult to see why not. So, it seems that a self-conscious subject need not ever meet any other

---

<sup>18</sup> See Evans, 1982, pp.103-104, and p.209.

people, just animals. And, following the rest of the argument through, we can then say that these other animals cannot be Cartesian egos, otherwise we would not be in a position to identify them. So far so good, but the point of this argument was, surely, to show that the self-conscious subject himself cannot be a Cartesian ego, and it is not clear how we are supposed to arrive at *that* conclusion.

This points us towards a serious defect with the argument, which can be pointed out in another way. For suppose the subject does make contact with other people. The argument, if sound, would show that *they* cannot be Cartesian egos. But from this claim, we cannot move to (5) which states that ‘if subjects are Cartesian egos, they cannot identify other subjects’. What it should say is that ‘if subjects are Cartesian egos, they cannot *be identified by* other subjects’. But a self-conscious subject, happily ascribing P-predicates to himself and to others, cannot infer from this proposition that *he* cannot be a Cartesian ego. To reach this conclusion, it would be necessary to add the premise that he and the others are of the same kind, or rather must be identified in the same way. But this, of course, is implausible. I identify others primarily by sight, but this not the way in which I primarily identify myself.

Furthermore, since we are interested in Strawson’s argument insofar as it lends support to the Body Claim, this consequence is disastrous. For, we want to get to the conclusion that any self-conscious subject must experience *himself* as embodied. But we have not even been given the materials to conclude that the subject must *be* embodied. There seems little hope of securing the further claim.

Let us suppose, however, that the proponent of the Body Claim can somehow diffuse this objection. There is another difficulty with the argument. For, we need to know how we are supposed to understand the word ‘identify’ as it occurs in steps (2)-(5). Strawson describes the notion of identification as follows,

It seems that the general requirements of hearer-identification could be regarded as fulfilled if the hearer knew that the particular being referred to was identical with some particular about which he know some individuating fact, or facts, other than the fact that it was the particular being referred to. To know an individuating fact about a particular is to know that such-and-such a thing is true of that particular and of no other particular whatever...This, then, is the general condition for hearer-identification in the non-demonstrative case; and it is obvious that, if a genuine reference is being made, the speaker, too, must satisfy a similar condition. (Strawson, 1959, p.23)

This description of identification is framed explicitly in terms of language. But Strawson sees no problem in transferring the basics of his account to the domain of thought. He claims that, “Each of us can *think* identifyingly about such particulars without talking about them”. (Strawson, 1959, pp.61).

Returning to the argument against dualism then, step (2) tells us that to ascribe a P-predicate to another subject, a subject must be able to identify that subject (with ‘identify’ understood in the way just described). A brief consideration of Strawson’s gloss on identification will show us that we should not accept this step. For Strawson’s general condition for identification is none other than Russell’s Principle. The condition for identification, or reference, is that the subject is able to distinguish the object of reference from all other things<sup>19</sup>. I argued in chapter three that this is too strong a condition on reference. If the argument given in that section against accepting Russell’s Principle is sound, then we should not accept step (2) of Strawson argument against dualism. We find that, once again, Russell’s Principle is being used to support an argument for the Body Claim<sup>20</sup>. However, as I have pointed out in previous chapters, Russell’s Principle is not something that the defender of the Body Claim can simply help themselves to.

It might be wondered whether Strawson’s argument could be reformulated without the aid of Russell’s Principle. It is difficult to say without knowing more about what notion of reference is in play. If, however, one were to endorse some form of causal account on which, to put it crudely, one can refer to an object if one is appropriately causally related to it, then it seems that we will have trouble supporting step (3). For, we no longer have to make sense of the idea that the subject knows some individuating description of the Cartesian ego in question<sup>21</sup>. It seems, then, that we cannot utilize Strawson’s argument against dualism in Chapter three of *Individuals* to support the Body Claim. With this negative conclusion I leave Strawson’s remarks on the conditions of self-consciousness. These considerations have not been entirely negative, however, since, as we shall see in chapter six, a line of thought can be

---

<sup>19</sup> In Part Two of *Individuals*, Strawson writes that, “in order for an identifying reference to be made, there must be some true empirical proposition known, in some not too exacting sense of this word, to the speaker, to the effect that there is just one particular which answers to a certain description.” (Strawson, 1959, p.183).

<sup>20</sup> I am not saying that Strawson was using it to support the Body Claim.

<sup>21</sup> Of course, this presumes that the dualist can solve the problem of mental causation. But, if successful, that problem scuppers Cartesian dualism all by itself. Strawson’s argument would be otiose.

developed which, whilst it owes much to Strawson's argument against dualism, rests on nothing so implausibly strong as Russell's Principle.

## 6. Self-Location

The self-location argument begins with self-reference. The starting assumption is that self-consciousness necessarily involves thinking about oneself. The argument proceeds as follows: First-person thoughts are thoughts about oneself. In thinking such a thought one thinks of oneself as a particular. Thinking of something as a particular involves thinking of it as in principle distinguishable from other particulars of the same sort. In order to think of oneself as in principle distinguishable from other things of the same sort, one must think of oneself as, in some sense, spatially located.

### 6.1 Thinking of Oneself as a Particular

The first step of the argument is the claim that first-person thoughts are thoughts about oneself as a particular. If one thinks that abstract objects are objects, then the class of particulars is a subclass of the class of objects. Thinking of *a* as a particular is to be distinguished from thinking of *a* as a property or as an abstract object<sup>1</sup>. Particulars might include physical objects like rocks and non-physical objects like souls. My claim is simply that in thinking the first-person thought ‘I am *F*’ I am thinking of myself as a particular in this sense. I will treat it as an assumption. I take it that the really controversial claim is that thinking of something as a particular involves thinking of it as in principle distinguishable from other particulars of the same sort. I will explain how I understand and support this claim.

Particular things have, what Evans calls, a fundamental ground of difference<sup>2</sup>. The fundamental ground of difference of an object at a time is that which differentiates it from other objects of the same kind. So, for instance, the fundamental ground of difference of a physical object at a time is its position in space (plus, possibly, its sort). As Evans puts it, the fundamental ground of difference of an object at a time will be an answer to the question ‘What differentiates that object from others?’, of the kind appropriate to objects of that sort.

As we have seen, Evans himself thought that to be able to think about an object, one must be able to distinguish it from all other things. This means that one

---

<sup>1</sup> I think it would probably also be reasonable to exclude thinking of oneself as an event.

<sup>2</sup> Evans, 1982, p.106-7. See chapter three.

must either be in possession of its fundamental ground of difference, or one must know what it would be for an identification of the object via its fundamental ground of difference, to be true. This is Russell's Principle. I argued in chapter three that Russell's Principle, as understood by Evans, is unacceptably strong.

Whether or not Evans' theory is ultimately acceptable, I think that the following aspect of it is surely correct: that to think of something as a particular (and hence as an object) one must think of it as having *some fundamental ground of difference or other*. That is, one must think of it as *in principle* distinguishable from other things of the same kind. To think of something without thinking of it as the possessor of a fundamental ground of difference at all, would not be thinking of it as an object (and hence a particular) at all.

What does this mean? It obviously must not mean that to think of something as an object one must have the concept of a fundamental ground of difference. Presumably, however, it *does* mean that one must have the concept of something being distinct from another. Thinking of something as an object requires that one think of it as distinct from all other things, even if it does not require knowing that in virtue of which it is so distinct.

Furthermore, it seems likely that in thinking of something as a particular one is not required to know what kind of thing it is. It seems intuitively correct to suggest that one can think of oneself even though one does not know, or truly believe, what kind of thing one is (physical object, non-physical soul etc.)<sup>3</sup>. Whilst I think that this is probably correct, it is controversial and I don't really require it for my purposes. What I need to claim is that in order to think of *a* as a particular one need not know what sort of fundamental ground of difference is relevant to objects of its kind. This falls short of the claim that in order to think of something as a particular I need not know what kind of thing it is. For, whilst I am pretty sure what sort of thing I am (a person), I am not entirely confident that I know what the relevant fundamental ground of difference for persons is (is it bodily or psychological?). Thus, my claim is that thinking of something as an particular means thinking of it as having some fundamental ground of difference or other, but one is not required either to know what that fundamental ground of difference is, or even what *kind* of fundamental ground of difference it has.

---

<sup>3</sup> This intuition did some work in Cassam's Identity Argument, see chapter four.

What *is* required, however, is that one think of it as in principle distinguishable from other things of the same kind. That is, to think of something as a particular, one must have the conceptual capacities to recognise that there could be two objects of that kind. This does not mean that one must have seen or otherwise experienced another object of that kind. Neither must one have even entertained the thought that there could be such a thing. All that is required is that one not be conceptually debarred from entertaining the thought that there could be another such object. The point is that thinking of something as having a fundamental ground of difference must be thinking of it in such a way as to leave room for the thought that there could be another such thing<sup>4</sup>. Unless one's way of thinking leaves room for this thought, one is not thinking of the object of one's thought as a particular at all.

I take these remarks to specify something like a necessary condition on thinking of something as a particular. It is, however, not a sufficient condition. Properties and abstract objects have fundamental grounds of difference also, and thinking of something as a property or an abstract object (arguably) requires that one think of it as the possessor of some fundamental ground of difference or other. But one is obviously not thereby thinking of it as a particular.

## **6.2 Thinking of Oneself as Spatially Located**

The next step in the self-location argument is the claim that in thinking of oneself as in principle distinguishable from other things of the same sort, one must think of oneself as, in some sense, spatially located. The idea is that without possessing spatial concepts and thinking of oneself as located in space, one does not have the resources, and so cannot be in a position, to think of oneself as a particular. The motivation for this claim is the well known problem that the Cartesian dualist faces in giving a coherent account of how it is that souls are individuated<sup>5</sup>.

Physical objects are individuated by their spatial location (plus, possibly, their sort). That is, the fundamental ground of difference at a time for a physical object is

---

<sup>4</sup> An objection: Surely we can think of the universe (everything that actually and possibly exists) as a particular, but such a thought does not leave room for the thought that there could be another such thing. My reply is that if my view entails that we cannot think of the universe (conceived in such a broad way) as a particular, this is an interesting consequence, rather than an objection.

<sup>5</sup> This problem is one of Strawson's favourite objections to dualism, see Strawson, 1959, Ch.3, 1966b, and Strawson's remarks at the end of Swinburne, 1987. Also see Sosa, 1987. It is at this point that the similarity between the present argument and Strawson's attack on dualism is most clear.

its location in space. This is a metaphysical point, it concerns what it is that makes it the case that there are two physical objects rather than one (or three rather than two). But there is an epistemological analogue. We can distinguish between distinct physical objects. In fact, if what I said in the previous section is correct, in order to think of something as a particular, one must be thinking of it as being in principle distinguishable from other things of the same kind. In the case of physical objects this constraint is easily met. In thinking of a physical object, I think of it as being located somewhere in space. This is not to say that I must know what its location actually is, just that I am thinking of it as having some location or other.

The difficulty arises for the dualist when we ask in virtue of what are non-physical souls individuated? What makes it the case that there are two souls rather than one? There are, as far as I can see, three answers the dualist could give to this question, each of which is problematic in its own way. The first would be to say that souls are individuated by the position in space of the bodies to which they are ‘attached’. The second would be to claim that souls are individuated by their position in logical space. The third would be to say that there is just a brute relation of distinctness which holds between souls and which cannot be further analysed<sup>6</sup>.

Now, the claim I am arguing for in this section is that without spatial concepts, and the ability to think of oneself as located, one could not think of oneself as a particular at all. So, let us imagine a subject with no spatial experience and no spatial concepts (supposedly) entertaining the first-personal thought ‘I am happy’. This subject, call him Vernon, must be thinking of himself in such a way that he is in principle distinguishable from other objects of the same kind (other subjects). By definition, Vernon cannot be thinking of himself as a physical object, since he has no spatial concepts. Similarly, he cannot be thinking of himself as a soul ‘attached’ to some physical object. Thus the first option that is open to the dualist is not an option of which Vernon can avail himself. So how is Vernon thinking of himself? Might Vernon be thinking of himself as individuated by his position in logical space? We can surmise that Vernon has certain beliefs (and other attitudes), and that these serve to pick out a ‘logical position’. Since two persons tend to have different overall belief structures, they will occupy different such positions. In this way, we might try to

---

<sup>6</sup> Swinburne’s dualism coupled with his claim that “personal identity is something ultimate, unanalysable” (Swinburne, 1984, p.26), suggests that he may hold such a view.

individuate souls. And some things (which may be objects of some sort) evidently *are* individuated in similar ways. For instance, Fregean thoughts are individuated by their contents, numbers are individuated by their position relative to the other numbers, and so on.

The problem with this idea is that it doesn't provide for qualitative identity with numerical diversity. If we individuate souls by their position in logical space it will follow that there simply could not be two souls which occupy the same logical location. It would be impossible for two subjects to have exactly the same overall belief structures. So it seems that this view is a non-starter as an account of the individuation of souls. Still, might Vernon still not *think* of himself in such a way? I find it difficult to form an opinion here. But I will say that *if* Vernon can think of himself in this way, and therefore as in principle distinguishable from different objects of the same sort, then Vernon is thinking of himself in a way which is radically different from the way in which we actually do think of ourselves. On this supposition, Vernon is thinking of himself as something akin to an abstract object. But the premise of the self-location argument is that self-consciousness involves thinking of oneself as a particular. If our premise is acceptable, Vernon cannot be self-conscious subject thinking of himself as an abstract object individuated by his position in logical space.

Could it be that Vernon is thinking of himself as something which could be distinguished from another by a brute, unanalysable, distinctness relation? I suggest that the answer to this should be no. The point of calling this distinctness relation 'brute' is that nothing more can be said about it. But surely unless one can offer an answer to the question 'In virtue of what are *a* and *b* different things?' then one cannot be thinking of them as distinguishable. It was agreed that to think of something as a particular one must not be conceptually debarred from entertaining the thought that there could be another such thing. But this suggestion endangers that claim. To think of something as in principle distinguishable from other things of the same kind, it seems that one must not be conceptually debarred from saying in virtue of *what* two such objects are distinct. The idea that two particulars, not thought of as located in space, might share all their properties and yet be distinct is not a thought that we can grasp. But this is precisely what the brute distinctness view must deny. The result, I think, is that Vernon cannot be thinking of himself as individuated by such a brute

distinctness relation<sup>7</sup>. This claim is closely related to that made by Strawson in the previous chapter, that subjects of experience must think of themselves as having empirically adequate criteria of identity.

My conclusion, then, is that to think of oneself as in principle distinguishable from other objects of the same kind, requires that one think of oneself as located spatially. But this is not an argument against dualism. I am not claiming either that one must be, or that one must think of oneself as, a physical object. I do not want to rule out the possibility that one could think of oneself as merely ‘attached’ to a located physical object; or as part physical object, part non-physical object; or as a geometrically located point of view. The point is that thinking first-person thoughts, thinking of oneself as a particular, requires that one have spatial concepts and that one use them in so thinking. The only way in which we could in principle distinguish between different subjects, is by their having diverse spatial locations.

Before I move on, I should say a few words concerning the relation between the arguments of the previous two sections, and some claims made by G. Strawson, 1999a, who claims that “many popular claims about the conditions of self-consciousness are too strong” (1999a, p.310). He defends the view that the notion of a self-conscious subject in the position I have assigned to Vernon is perfectly coherent. Obviously I disagree. I want only to point out that what one might think are two of the most controversial claims I make during the course of my argument he actually endorses. These are the claims that first-person thought is thought about oneself as a particular (1999a, p.314), and that a first-personal thinker must not be conceptually debarred from forming and entertaining thoughts of a subject of experience other than himself (1999a, p.310). I thus claim to be giving an argument which contradicts G. Strawson’s view, using materials that he himself accepts.

### **8.3 Spatial Concepts and Spatial Experience**

The self-location argument concludes that a self-conscious subject must think of him or herself as spatially located, and thus be in possession of spatial concepts. Both the

---

<sup>7</sup> This view is defended in Sosa, 1987. He writes, “entities x and y *cannot possibly be related simply by diversity*... The problem with supposed nonspatial particulars like souls is simply that no pure relation can be found to play the role that spatial apartness plays for spatial particulars. We have no inkling of any relation that can help us grasp the diversity of two souls alike in every qualitative respect” (Sosa, 1987, p.69).

solidity and the action arguments, however, start from the premise that a self-conscious subject must enjoy spatial *experience*. Can we forge the necessary tight connection between spatial concepts and spatial experience?

It seems clear that we will not get a firm answer to this question by perfectly general considerations on the nature of concepts. One might think that a thoroughgoing empiricism or maybe some defence of the claim that ‘concepts without intuitions are empty’ might help us here. But it is entirely reasonable to point out that there are a range of different kinds of concepts. On the one hand we have observational concepts such as *red*. It would not be surprising to find that the concept *red* cannot be grasped by a subject without the subject having undergone some experiences as of red things. Call these concepts experience-dependent. On the other hand, we have theoretical concepts such as *electron*. No-one has ever experienced an electron (except indirectly), or had any experience as of an electron, and yet it is a concept which many of us have. Call these concepts experience-independent. The questions we need to answer in the present context are these: What are the spatial concepts that the self-location argument shows are required for self-consciousness? And, are these concepts experience-dependent or experience-independent?

Here is a simple argument for the claim that the concepts involved are experience-dependent. The self-location argument concludes that, if one is self-conscious, one must be able to think of oneself as spatially located. That is, one must be able to think, ‘there is some place where I am’. We can call this place *P*. *P* is either a descriptive name<sup>8</sup>, or an abbreviation of a definite description. The description in question being ‘the place where I am’. Thus, one must be able to think, ‘I am at *P*’. But, the argument continues, there is no difference between thinking ‘I am at *P*’ and thinking ‘*P* is here’, or even ‘I am here’. The reason for this is that ‘*P* = here’ is surely uninformative, as uninformative as is ‘the place where I am = here’<sup>9</sup>.

If this is correct, then a self-conscious subject must be able to think ‘here’-thoughts. Now, thinking a ‘here’-thought means thinking a thought that refers to the place at which the thinker is. But, for such a ‘here’-thought to refer, the subject must know which place it is to which he or she is referring, or, at least, he or she must be disposed to find out which place it is. This means that a self-conscious subject must

---

<sup>8</sup> See Evans, 1982, §1.7.

<sup>9</sup> This is why ‘Here’ is such an unhelpful answer to the question ‘Where are we?’.

be, at least disposed to, experience the place to which he or she refers with ‘here’. So, a self-conscious subject must have some kind of spatial experience.

It should be immediately clear what is wrong with this argument. The argument relies entirely on Russell’s Principle<sup>10</sup>. Having rejected Russell’s Principle in the case of ‘I’-thoughts it would be strange to offer such strikingly different account of ‘here’-thoughts. Rather, we should agree that ‘here’ refers to the place of utterance. Furthermore, in thinking a ‘here’ thought I am thinking of a place as a place, but this does not require me to be able to distinguish that place from all other places.

Maybe I must be disposed to treat any information I receive from the place at which I am as germane to my ‘here’-thoughts, and such information might also dispose me to act in relevant ways<sup>11</sup>. But it would take a further argument to show that these dispositions must become manifest at some point. So this point should not be taken to imply that I must at some time actually receive some information from, or have experience of, the place where I am. This line of thought is rejected by Evans, who writes that

Where there is no *possibility* of action and perception, ‘here’-thoughts cannot get a grip. Consider the philosopher’s fantasy of a brain in a vat: a person’s thinking organ kept alive and capable of sustaining thoughts, yet with no avenue of perception or mode of action. We can perhaps imagine being the person whose thoughts are sustained in this way; but in the perceptual darkness and silence of our existence, we could surely have no use for ‘here’. If we knew what had become of us, we could certainly think of a place as the place where the brain which sustains our thoughts is located – but this is a mode of identification of a place quite unlike that expressed by ‘here’. (Evans, 1982, p.153)

If the self-location argument is acceptable, then the brain in the vat must be thinking of himself in such as way as to make room for the thought that he is located somewhere in space. So, he must be able to think that there is some place where he is. The argument above suggested that this is equivalent to saying that there is some place where the brain in a vat can think of as ‘here’. But even if Evans is right, and this mode of thinking of a place is unlike the way in which *we* think of places as here,

---

<sup>10</sup> Actually, it relies on a ‘knowing which’ condition on reference. And it might be that an account of ‘knowing which’ can be given which falls short of Evans’ ‘discriminating knowledge’. If such an account can be given, there may be a more direct route than that presented from spatial concepts to spatial experience.

<sup>11</sup> These aspects of ‘here’-thoughts are central to Evans’ account, see Evans, 1982, §6.3.

why should this matter? All that is required is that the brain in the vat can have a thought about a place, and this does not seem to be in question.

But this is not the end of the matter. For the self-location argument establishes that a self-conscious subject must be thinking of himself as spatially located. But thinking of oneself as spatially located is not just a matter of thinking a thought which is, as a matter of fact, about a place. If one's thought is to count as genuinely spatial, one must be thinking about a place *as* a place (or about a spatial property *as* a spatial property). And this means that one must have an understanding of what a place is. And now the thought is that it is highly plausible to suggest that one cannot have a full grasp of the notion of a place unless one has had some experience as of places. A subject who has had absolutely no experience as of places cannot really be said to know what places are. This seems right *even if* such a subject can entertain thoughts which are about places. A reason for thinking that this is correct is that it is difficult to imagine being able to explain what places are to a subject who lacked any spatial experience. In this respect 'here' is more like 'red' than 'electron'. To repeat, thinking about what is in fact a place is not sufficient for one's thought to be genuinely spatial. One must be thinking of a place *as a place*. And it is this which only seems possible on the assumption that the subject has had some kind of spatial experience.

If this is correct, then the self-location argument establishes more than just that a self-conscious subject must be capable of thinking of himself as spatially located. It actually establishes the further conclusion that a self-conscious subject must have, or have had, experience which counts as genuinely spatial. That is, he must have had experience of places *as* places (or spatial properties *as* spatial properties). This is important, since both the solidity argument of chapter seven and the action argument of chapter eight begin from the assumption that self-consciousness requires spatial experience. Both of these arguments claim that bodily-awareness is a transcendental condition of spatial experience, and it is to these arguments that I now turn<sup>12</sup>.

---

<sup>12</sup> It is plausible to think that once we have established that a self-conscious subject must enjoy spatial experience, we can establish the objectivity condition (both OC1 and OC2). Although I will not attempt such a feat, it seems reasonable to think that in order to perceive spatially one must be able to employ both the seems/is distinction and the self/world distinction. Evans writes, "The idea that there is an objective world and the idea that the subject is somewhere in the world cannot be separated" (Evans, 1982, p.222).

## 7. Solidity

The solidity argument takes as its starting point the assumption that self-consciousness requires spatial experience. Taken independently of the self-location argument it can be thought of as an attempt to spell out the transcendental conditions of self-conscious subjects who, like us, have some kind of spatial experience. The idea driving the argument is that spatial experience is necessarily tied up with the experience of solidity, and that the experience of solidity is something which requires bodily-awareness<sup>1</sup>.

### 7.1 Spatial Experience and Physical Objects

The solidity argument can be stated as follows:

- (1) Spatial experience requires experience as of physical objects.
- (2) Experience as of physical objects is only possible if one has experience as of solidity.
- (4) Experience as of solidity is only possible for a subject with a sense of touch.
- (5) Touch requires bodily-awareness.

So,

- (6) Spatial experience is only possible for a subject with bodily-awareness.

A subject with bodily-awareness is clearly a subject who satisfies Cassam's 'second grade of apparent presence in the world', i.e. who experiences himself as having a body. Therefore, if the solidity argument is sound, the weak version of the Body Claim is true.

Let us begin with the first premise. A line of thought which can lead one to accept this claim is to be found in Ch. 1 of Strawson's *Individuals*. The idea is that,

if we are to operate the scheme of a single unified spatio-temporal system or framework of particulars...we must have criteria or methods of identifying a particular encountered on one occasion,

---

<sup>1</sup> Arguments which resemble the solidity argument, as presented below, are given in Joske, 1967, Ch.2, O'Shaughnessy, 2000, Ch. 23, and Cassam, 2002.

or described in respect of one occasion, as *the same individual* as a particular encountered on another occasion, or described in respect of another occasion...The necessity may be brought out in the following way. It is not the only way. Evidently we can sometimes referentially identify a member of the spatio-temporal framework by giving, or being given its position relative to others. No less evidently we cannot make the identification of every element in the system in this way relative to that of other elements. An immediate answer is that we have no need to, because we can identify some elements by direct location. But this answer, by itself, is insufficient. For we do not use a different scheme, a different framework, on each occasion. It is the essence of the matter that we use the same framework on different occasions...We cannot attach one occasion to another unless, from occasion to occasion, we can reidentify elements common to different occasions. (Strawson, 1959, pp.31-2)

To summarise, our spatial experience is an experience of a single unified spatio-temporal system. But in order for this to be the case, different 'occasions' must be 'attached' to one another. That is, each stretch of continuous observation must be attached to another. But this is possible only on the assumption that we can sometimes regard two experiences as experiences of one and the same thing. Therefore, spatial experience is necessarily such as to support the possibility of the reidentification of particulars.

Why does the possibility of the reidentification of particulars require experience as of physical objects? To be able to attach together different stretches of continuous observation, I must be able to think of two experiences of places as experiences of the same place. We can, for the time being, take this much as having been established. The important claim, now, is that the ability to do this presupposes the ability to reidentify not only places, but physical objects. This dependence goes in both directions,

the identification and distinction of places turns on the identification and distinction of things; and the identification and distinction of things turns on the identification and distinction of places (Strawson, 1959, p.37)

For instance, we know that we are once again in the same place, say Trafalgar Square, because we know that the object we now perceive is the very same object that we perceived when last in Trafalgar Square, say Nelson's Column. And conversely, we know that we are seeing Nelson's Column, because we know that we are once again

in Trafalgar Square<sup>2</sup>. All this means that if we had no experience as of physical objects, we would have no way to reidentify places, and thus no way in which to attach together different stretches of continuous observation.

There are, at least, two objections to this argument. First, it will be pointed out that although *our* spatial experience is of a single unified spatio-temporal system, we have not been shown why this is necessarily so. And it is this latter claim which is needed if the solidity argument is to be genuinely transcendental. Indeed, this latter claim is not at all obvious, since we appear to be able to imagine experience as of two spatially unrelated spaces<sup>3</sup>. If we can do this, it seems that we can imagine experience as of many spatially unrelated, but qualitatively identical spaces (maybe changing each time we blink). There is no reason to think that such an experience would not be spatial. And it would appear to satisfy the demands of the self-location argument – a subject in this situation would be able to locate themselves at any given time. The very first step in this transcendental argument fails.

A reply to this objection might proceed as follows<sup>4</sup>: the subject of such experience must think of him or herself as unified and, in particular, causally connected over time. But, this amounts to thinking that the spaces experienced are unified and causally connected. The unity of the subject requires the unity of the spaces the subject experiences. Thus, the spaces are not experienced as entirely spatially unrelated. But, in the present context, this is unconvincing. First, whilst it may be true that one must think of the spaces one experiences as causally connected, it does not automatically follow that one must think of them as spatially related. It may not be possible, but science fiction seems to have got us used to the idea of causal relations in the absence of spatial relations. The reasons why the unity of the subject might be thought to require the unity of the spaces experienced is that the subject is thought of as tracing a continuous path over time. But we need only think of the subject in this way if we think of it as a physical object, or at least behaving in the same way as physical objects behave. The point is that at this stage in the argument, this is not something that we are allowed to assume.

---

<sup>2</sup> Williams, 1961, p.108, suggests that this amounts to accepting a relational view of space. Even if Strawson does have something like this in mind, which seems somewhat doubtful, nothing I say commits me to such a view.

<sup>3</sup> Maybe as in Quinton, 1962.

<sup>4</sup> Eilan, 1993, puts this line of thought to a slightly different use.

It seems to me, then, that this is a perfectly valid objection<sup>5</sup>. And this is a point at which we must add in an empirical assumption if we are to give the solidity argument the chance to fly. The assumption we must add is that the self-conscious subjects which we are discussing are such as to enjoy experience as of a single unified spatio-temporal system. And this is not such a huge problem. For, even with this assumption made, we might still come up with some philosophically interesting claims. If it turns out, as the solidity argument claims, that it is a necessary condition of self-consciousness with experience as of a single unified spatio-temporal system, that one experience oneself as having a body, then we will have established something of real interest. So, even though we do not seem to be able to rule out a priori the possibility of a self-conscious subject who does not enjoy experience of a single unified spatio-temporal system, this should not worry us to much.

Another objection has been raised to this argument by Campbell<sup>6</sup>. Campbell argues that the reidentification of (and thus experience as of) physical objects is not required for the reidentification of places. For places can be reidentified on the basis of the reidentification of ‘located features’. Features are something less than physical objects, the relevant difference being that physical objects, but not features, display ‘internal causal connectedness’. So,

In contrast to features, the condition of a thing at any one time is thought of as being causally dependent upon its condition at earlier times. One of the determinants of its properties at a given time is which properties it had earlier, and this is so no matter how much it has moved around. It might, indeed, be held that internal causal connectedness is more basic than spatio-temporal continuity, in our notion of objecthood... We can certainly make sense of someone finding their way around using stably located features, or groups of features, which they do not think of having that internal causal connectedness over time (Campbell, 1993, p.67)

If place reidentification is possible on the basis of feature reidentification, and feature reidentification does not require experience as of physical objects, this objection is a serious one. But let us look more closely at this supposed reidentification of features. Clearly *we* can reidentify places by reidentifying features. For instance, I can reidentify a place by “the pool of light thrown by a projector onto a wall” (Campbell, 1994b, p.8). But this is not yet to show that a subject who has no experience

---

<sup>5</sup> But see Strawson’s reaction to this kind of worry, 1959, pp.32-6.

<sup>6</sup> See Campbell, 1993, 1994a, and, 1994b.

whatsoever as of physical objects could reidentify places purely by reidentifying features. A subject who only has experience as of features *ex hypothesi* does not think of their present state or location as in any way dependent upon their former state or location. But how, then, can such a subject think of a currently perceived feature as *the very same* feature as was perceived before? This subject would have no reason to think of the currently perceived feature as numerically, rather than just qualitatively identical. But it is numerical identity that is required for reidentification.

Of course, for us, there is a distinction between the qualitative and numerical identity of features. But it is plausible to think that this is a result of our thinking of them as, in some way, dependent upon physical objects. Take the pool of light thrown out by the projector. What licenses us to treat a currently perceived pool of light as numerically rather than just qualitatively identical to a remembered one? Arguably, it is the knowledge that it is being thrown out by the very same projector and that the projector has not been switched off in the meantime<sup>7</sup>. And the projector is something that is being thought of as internally causally connected, and thus as more than just a feature.

My suggestion, then, is that a subject for whom everything is perceived as a feature is not in a position to grasp the distinction between qualitative and numerical identity. And this is a subject for whom there is no possibility of reidentification of features, and correspondingly of places. So, I think, Campbell's objection is unsuccessful. We can, then, provisionally accept that spatial experience requires experience as of physical objects.

## 7.2 Solidity and Force

The second step in the solidity argument is that experience as of physical objects is only possible if one has experience as of solidity. An argument for this claim can be formulated in, roughly, the following way: physical objects are essentially solid therefore, to experience something as a physical object is to experience it as solid. A proper assessment of this argument requires us to first of all say something about the notion of solidity.

---

<sup>7</sup> An alternative answer might be, 'that there is no evidence to the contrary'. But what could possibly count as evidence to the contrary for a subject who did not have the concept of a physical object?

Let me suggest three uses of the word ‘solid’ (there are undoubtedly more). First, ‘solid’ may be contrasted with ‘liquid’ and ‘gas’, these terms denote three distinct states of matter in physics. Second, ‘solid’ is often used as roughly interchangeable with ‘hard’, as when one says, ‘this chair is certainly a solid one’. Third, ‘solid’ can be used to mean something that is closely related to *impenetrable*, meaning that solid things cannot penetrate or be penetrated by each other<sup>8</sup>.

Taken in either of the first two senses, it is clearly not the case that experience as of physical objects requires experience as of solidity. A ball of mercury is experienced as a physical object even though it is experienced as a liquid, and so not solid in the first sense. Again, a jelly is experienced as a physical object even though it is obviously not experienced as solid in the second sense. If we are to lend plausibility to this step of the solidity argument, then, I think it would be useful to turn to the understanding of solidity in terms of impenetrability.

Solidity is listed by Locke as a primary quality<sup>9</sup>, but what *is* solidity? There is clearly a very close connection between solidity and impenetrability but Locke suggests keeping the two notions separate,

That which thus hinders the approach of two bodies, when they are moving one towards another, I call *solidity*...if anyone think it better to call it *impenetrability*, he has my consent. Only I have thought the term *solidity* the more proper to express this *idea*...because it carries something more of positive in it than *impenetrability*; which is negative, and is perhaps more a consequence of *solidity*, than *solidity* itself. (Locke 1975, Bk.II, Ch.IV, §1)

Locke has more to say about solidity,

This, of all others, seems the *idea* most intimately connected with and essential to body...[it is] inseparably inherent in matter...This is the *idea* which belongs to body, whereby we conceive it *to fill space*. (Locke 1975, Bk.II, Ch.IV, §1-2)

Here Locke claims that all material stuff has the property of solidity, and further that all physical bodies/objects have the very same property (remember, not all material stuffs are objects). What’s more, the solidity of physical objects might well be said to constitute their *essence*. Solidity (space-fillingness) is what is essential to the

---

<sup>8</sup> This is how Locke, 1975, Bk.II, Ch.IV, views solidity.

<sup>9</sup> Locke, 1975, Bk.II, Ch.VIII, §9.

physical, and it gives rise to impenetrability (the power to entirely exclude other space-filling objects)<sup>10</sup>.

Is it correct to say that physical objects are absolutely impenetrable, i.e. have a property which, “excludes all other solid substances” (Locke 1975, Bk.II, Ch.IV, §2)? Two objections to this claim present themselves, one easily dealt with, the other less obviously so. The first is that such things as swarms of bees, or clouds, by no means exclude other such things. Two swarms of bees, or clouds, may pass through each other. There are a variety of responses to this objection. First, it might be held that, although swarms of bees and clouds are composed of physical objects, they are not themselves physical objects. Bees and water particles are physical objects, swarms and clouds are not. Second, since no two bees or water particles occupy exactly the same space, the swarms, or clouds, are not in exactly the same places. Each resides within the other’s gaps. Third, it could be argued that the two swarms, or clouds, momentarily become one swarm. Since any one of these replies seems unproblematic, the objection is unconvincing.

The second objection to the claim that physical objects are necessarily impenetrable is that, contrary to the present hypothesis, it is perfectly possible for two solid objects to be in the same place at the same time. This will be the case with statues and lumps of bronze, and maybe even people and their bodies<sup>11</sup>. Objects may very well be solid, in some sense, but they are not impenetrable. Thus, whatever solidity is, it cannot be the same as absolute impenetrability. Some will argue that this objection is misconceived, and that the language of ‘material constitution’ is not the right way to think about physical objects<sup>12</sup>. But, at the very least, we should be worried about equating the solidity of physical objects with the impossibility of two such objects occupying the same place at the same time<sup>13</sup>.

---

<sup>10</sup> A nice discussion of Locke’s notion of solidity can be found in Alexander, 1991. Locke’s views on solidity are vigorously disputed in Hume, 1978, Bk.I, Pt.IV, §4. That all physical objects are impenetrable is accepted by Quinton, 1964. Impenetrability also makes it on to Bermudez’s list of ‘object properties’, 1998, pp.72-73, although it is unclear how this relates to a specifically Lockean understanding.

<sup>11</sup> See Wiggins, 1968.

<sup>12</sup> See, for example, Burke, 1994.

<sup>13</sup> It is true that, in his discussion of impenetrability, Locke seems to be making the point that body’s being impenetrable just amounts to the claim that one cannot simply *squeeze* it out of existence. Or, rather, that one cannot *compress* the matter of one object into exactly the same place as that of another. Does the example of material composition not somehow bypass this aspect of Lockean solidity? Possibly, but I think that the language of compression is inessential. Lockean solidity prohibits that passing of one object *through* another. But thought experiments concerning memory-transfers and so on may well lead us to think that a person could pass *through* a body.

The way forward, I suggest, is to weaken the notion of solidity that is in play. We need an essential property of physical objects which falls short of solidity conceived as impenetrability. We might turn, in this vein, to Peacocke's, 1993, discussion of the property of being subject to mechanical forces,

I suggest that for something to be a quantity of matter is for changes in its state of motion to be explicable by the mechanical forces acting on it, and for its changes in motion to exert such forces (Peacocke, 1993, p.170)

If this is correct, then since physical objects are quantities of matter, physical objects are essentially subject to mechanical forces. Peacocke treats this claim as an assumption. As there seem to be no obvious counterexamples (physical objects not subject to mechanical forces), I will follow his lead. It looks as though we are now in a position to claim that experience as of physical objects requires experience as of force. For, since the notion of force is so 'intimately connected' with that of a physical object, if we experience something which we do not experience as solid, we are not experiencing it as a physical object at all. Thus, we turn the solidity argument into the mechanical forces argument.

But we have moved rather quickly here. First, it is not clear why the fact that physical objects are necessarily subject to force means that to experience something as a physical object is to experience it as being subject to force. Second, it might be held that all that is required to experience something as subject to force is that it be *thought of* as being subject to force. But this need not require that one has experience as of force. I take each of these difficulties in turn.

Let us suppose, for a moment, that the laws of nature are necessary<sup>14</sup>. If so, it will be the case that physical objects are necessarily subject to Newton's law of cooling, which states that the rate of cooling of a hot body which is losing heat both by radiation and by natural convection is proportional to the difference in temperature between it and its surroundings<sup>15</sup>. But, we should not infer from this that in order to experience something as a physical object one must experience it as subject to Newton's law of cooling. Since this is the case, we should not move from the fact that physical objects are necessarily subject to mechanical forces, to the conclusion that to

---

<sup>14</sup> As is argued in Shoemaker, 1998.

<sup>15</sup> Ignore the fact that this law does not hold for large temperature excesses.

have experience as of physical objects, one must have experience as of things which are subject to mechanical forces. Even if physical objects are necessarily *F*, it does not follow that to have experience as of physical objects one must have experience as of *Fness*. This runs counter to Peacocke's own view. He claims that the following is a "general principle governing the ascription of content" (Peacocke, 1993, p.169),

If an account of what is necessarily involved in something's having a certain property makes reference to some substantial condition which must be met by things which have it, a thinker's mental representations of that property must be suitably sensitive to the existence of this substantial condition. (Peacocke, 1993, 171)

The relevant property is that of being a physical object, and the substantial condition is that of being subject to mechanical forces. If what I have said about Newton's law of cooling is correct, this principle is not entirely general. However, the particular application of it to the substantial condition of being subject to mechanical forces does command some intuitive plausibility. As Peacocke writes,

Consider a case in which (say) a creature is in fact surrounded by holograms, which come into 'contact' with one another... This subject is unsurprised when these objects exert no force on him when they 'touch' his surface, and does not expect to have to exert any force to 'move' one of them out of his way. I have no objection to saying that this subject conceives of these things in his environment as objects. But... I do not think that he conceives of them as material objects. (Peacocke 1993, p.174)

What presumably explains the difference in our intuitions about the two different substantial conditions is that being subject to mechanical forces is far more central to our concept of a physical object than is being subject to Newton's law of cooling. Since being subject to mechanical forces does appear to be central to our concept of a physical object, the relevant instance of Peacocke's application seems correct<sup>16</sup>. It is no objection to this that there *could* be counter instances to the principle.

The second objection maintains that all that is required to have experience as of physical objects is that one *think* of them as being subject to mechanical force. One

---

<sup>16</sup> Other substantial conditions central to our concept of a physical object include being internally causally connected, and being an occupier of space. I would claim that Peacocke's principle holds true of these substantial conditions also. That is, unless one has experience as of internally causally connected space occupiers (that are subject to mechanical forces), one does not have experience as of physical objects.

need not experience them as such. It is clearly the case that we experience some things as physical objects without ever experiencing them as being subject to mechanical force. Maybe this is the case with the moon (for all but those lucky few who have been there). But for this to be an objection to the relevant step of the solidity argument, it must be maintained that a subject could have experience as of physical objects without *ever* having experience as of mechanical force. Consider a subject whose experience consists entirely in visually perceiving two moving spheres which never come into contact with each other. This subject has no experience whatsoever of mechanical force. The objection is that he might still be thinking of the spheres as subject to mechanical force, and thus able to experience them as physical objects. He might think that *if* they were ever to touch each other, then they would resist each other's movement, thus his mental representations are 'suitably sensitive'.

One way to counter this objection would be to claim that thinking of an object as subject to mechanical force is only possible if one has had some experience as of mechanical force. In the terminology of chapter six, this is the claim that the concept of mechanical force is experience-dependent. Earlier I claimed that it is plausible to think that some spatial concepts are, in an important sense, experience-dependent. But can we say the same thing about the concept of mechanical force? Certain remarks made by Peacocke, and expanded upon by Cassam, suggest that the answer to this question should be yes. Peacocke writes that,

someone capable of experiences of force may have available ways of thinking of that magnitude which are unavailable to someone who does not have experiences of the same general kind. (Peacocke, 1993, p.173)

Cassam concurs,

like other concepts of primary qualities, these are concepts which someone who has them must be in a position to apply on the basis of experience...A thinker who has never had any experience of shape or force is one who has no proper conception of shape or force. (Cassam, 2002, pp.328-329)

Unfortunately, we are offered little by way of argument for this claim. Maybe the claim does have some intuitive force, but it seems a great deal less convincing than the claim made in chapter six about the experience-dependence of spatial concepts such as *here*. In that case, we were considering a subject with no spatial experience

and asking whether he or she could be in possession of spatial concepts. As such, that hypothetical subject was in a far more attenuated experiential situation. Lacking all spatial experience, it seemed plausible to think that he or she lacked spatial thought. But the present hypothetical subject is in an experientially richer situation. True, there is no experience as of mechanical forces, but we as yet have no reason to think that the content of the relevant visual experience would differ in any significant way from ours. Removing all spatial experience seems more likely to deprive us of ways of thinking than does removing all experience as of mechanical force.

Of course, these remarks are highly speculative. It might turn out that this apparent gap in the solidity argument can be filled by some plausible general principle governing the ascription of conceptual capacities. It is advisable, then, to look at the remainder of the solidity argument to see whether or not it is worth searching for such a principle.

### **7.3 Touch and Bodily-Awareness**

What are the transcendental conditions of the experience as of mechanical force? There are two steps to be looked at in this section. One is the claim that experience as of mechanical force necessarily requires the sense of touch. The second is that the sense of touch comes hand in hand with bodily-awareness. Together, these amount to the claim that only a subject with bodily-awareness could experience objects as being subject to mechanical forces. For ease of presentation it will be useful to take the two claims in reverse order.

So, what is the relationship between touch and bodily-awareness? Following O'Shaughnessy, 2000<sup>17</sup>, we can distinguish between six principle types of tactile perception: momentary point contact, extensive contact, passive touch involving movement, guided but passive touch involving movement, active exploratory movement, and tactilely led circumambulation. O'Shaughnessy makes three major claims concerning these types of tactile perception, two of which I take to be rather implausible. The first is that the complex cases are 'built out of' the simple case of momentary point contact. The second is that it is not until limb-movement is

---

<sup>17</sup> The relevant chapters (23 and 24) of this book follow, almost word for word, his two earlier papers, 1995, and, 1989. I refer to the more recent version since, where small changes have been made, I assume this represents O'Shaughnessy's current view.

introduced into the tactile project that we are able to perceive the shapes of objects (according to O'Shaughnessy, this occurs to a limited extent in guided but passive touch involving movement, but is only fully apparent in active exploratory movement). Finally, O'Shaughnessy claims that the sense of touch is dependent on bodily-awareness.

The implausibility of the first claim can be seen by considering the simple case. The information gained in momentary point contact is 'something solid here' (O'Shaughnessy, 2000, p.662). One of the interesting features of this simple case is how the 'here' is specified. It seems to be something like end-of-finger-attached-to-hand-attached-to-arm-etc-stretched-out-in-front-of-body. The 'here' is specified as a body-relative point and the information includes both the point on a particular limb and the general orientation of the entire body. The properties of the object which are disclosed are *solidity* and *here* (understood as body-relative position). In line with the previous discussion, this solidity should not be conceived as 'absolute impenetrability', but as 'subject to, and exerting mechanical force'. I will adopt this usage for the remainder of the chapter. In the awareness of the 'here' aspect, I am aware of the orientation of my own body through bodily-awareness. My attention can be directed either inward towards my body or outward towards the object. In fact it seems plausible to say that my attention is always both inward *and* outward, one focally and one peripherally. When I switch perspective from inward to outward (or *vice versa*) I am shifting the balance between focal and peripheral awareness (Bermúdez, 1998, pp.138-9)<sup>18</sup>. In O'Shaughnessy's opinion this simple case provides the material out of which the complex cases of tactile perception are constructed. So, for instance, active exploratory movement is built out of lots of related 'something solid here's.

There are several reasons for thinking that the simple case fails as an adequate building block for complex perceptions. First, the idea that solidity is perceived in momentary point contact is highly implausible. The solidity of an object is only really discoverable given the action or movement of the perceiver<sup>19</sup>, for we perceive that an

---

<sup>18</sup> In his classic text on touch, the psychologist David Katz writes, "A subjective component which refers to the body seems inescapably linked with a second component that refers to the properties of objects. We therefore describe tactual phenomena as *bipolar*...At any moment either the subjective or objective side of tactual perception may be dominant, but this bipolarity nevertheless exists" (Katz, 1989, p.41).

<sup>19</sup> Or, possibly, the movement of the object one is in contact with.

object is solid through its offering resistance to movement<sup>20</sup>. Even if this claim is rejected it is certainly convincing to think that the perception of something as solid requires the *possibility* of movement, and that this requires the temporal dimension lacking in momentary point contact. So, maybe the information of point contact is just ‘some sensation here’. But if this is the case, it becomes difficult to see how the information given in complex tactile perceptions, which includes more than just sensations, can be built out of the purely sensational simple case.

The second reason for suspicion is a general mistrust of such atomistic accounts of perceptual experience. The idea that the perception of an object is composed of lots of perceptions, one of each of its parts, just grossly misrepresents the phenomenology of perception, tactile or otherwise. Perceptual experience doesn’t seem to be like that, so we would need good reason to believe that the structure of tactile perception is actually of this form. As far as I can see, O’Shaughnessy gives us no such reason.

What about the second claim, that it is not until limb-movement is introduced into the tactile project that we are able to perceive the shapes of objects? O’Shaughnessy’s claim is that neither momentary point contact nor extensive contact provide us with any determinate information concerning the shape of the object perceived. Such shape properties are only perceptible once the element of limb-movement has been introduced (O’Shaughnessy, 2000, p.665). So, resting my hand against an object does not reveal the shape properties of that object to me. On the other hand, if someone guides my passive hand around the edges of the same object I *will* gain that information (even more so if I actively feel the object). Thus, O’Shaughnessy makes limb-movement a condition of the tactile perception of shape properties. But this seems false. Extensive contact is not so poor in spatial determinations as O’Shaughnessy would have us believe. Rather, we are to some extent able to discern the shapes of objects in this way. For instance, Martin’s suggestion that we can imagine someone who could tell the difference between a 50-

---

<sup>20</sup> This claim has been disputed by Fricke and Snowdon, forthcoming. They point to experiences such as moving down a slide in which solidity is perceived but not as offering resistance to the movement of the perceiver, quite the opposite. This is, in a sense, correct but it is not damaging to the claimed relation between solidity and resistance. Of course, the solidity of the slide *facilitates* the movement of the perceiver, but it is also experienced as preventing the perceiver from falling immediately downwards. As Locke says, “Whether we move, or rest, in what posture soever we are, we always feel something under us, that supports us, and hinders our farther sinking downwards.” (Locke, 1975, Bk.II, Ch.IV, §1). This is no less true when one is, for instance, on a slide.

pence and a 10-pence piece by having them pressed against their skin, seems perfectly reasonable<sup>21</sup>. The point remains, however, that shape properties are *much easier* to perceive tactilely when limb-movement is introduced. It is just that some shape properties can be perceived with stationary limbs. As I suggested above it is perception of the property of *solidity* that requires some movement (or at least the potential for movement) on the behalf of the perceiver.

These considerations highlight the major divergences of O'Shaughnessy's and my own views of touch and its phenomenology. This said, I think that the most important aspect of O'Shaughnessy's account of touch (and its relation to bodily-awareness) is substantially correct. His idea is that we come to know the spatial properties of tactilely perceived objects through coming to know the spatial properties of the bodily movements involved in the perceptual project. Thus, in (active) touch, we know the perceived object to be round because we know that our hand has traced a circle<sup>22</sup>. There is a match between the spatial property of the perceived object and of the perceiving bodily movement. The awareness we have of the spatial properties of our bodily movements is not *prior* to the awareness we have of the spatial properties of the objects perceived. On the contrary, for the most part our attention is directed outwards towards the object. Nevertheless, an awareness of the bodily movements of the perceiver is a necessary *mediator* in tactile perception. Since this awareness of one's own bodily movements etc. is an exercise in bodily-awareness, this shows an important sense in which tactile-sense depends on bodily-awareness<sup>23</sup>.

I suggest that we accept this third claim of O'Shaughnessy's. That is, we should accept that without an awareness of the spatial properties of body parts, bodily movements, or bodily sensations, there would be no tactile awareness of objects. That this is the case can be seen from the preceding considerations. For the content of, both simple and complex, tactual perception includes a body-relative location. Awareness

---

<sup>21</sup> Martin, 1992, p.208.

<sup>22</sup> O'Shaughnessy, 2000, p.676.

<sup>23</sup> This claim has recently been rejected by Scott, 2001. Scott offers three groups of counterexamples to O'Shaughnessy's account of this close relation between the content of tactual perception and the content of bodily-awareness. These are; the perception of non-spatial properties such as oiliness and stickiness, dynamic tactile perception with held objects such as pencils or metal rods, and certain tactile illusions. Scott claims that in each of these cases, there can be tactile perception without the corresponding state of bodily-awareness. However, I don't think that these supposed counterexamples are particularly damaging to the solidity argument, for even if Scott has managed to come up with individual counterexamples, he has not shown (nor, I imagine, would he wish to claim) that a subject could enjoy tactual perception whilst entirely lacking any kind of bodily-awareness.

of sensations or felt properties at a body-relative location is an exercise in bodily-awareness. Furthermore, the tactile perception of all but the most basic of shape properties is only perceptible given bodily-movement. We know that the object is round because we know we have traced a circle. We know *that* through bodily-awareness. So, touch is dependent on bodily-awareness.

We now come to the other (postponed) step in the solidity argument. This is the contention that solidity (being subject to, and exerting mechanical forces) is only perceptible to a creature endowed with a sense of touch. We have seen that solidity *is* perceptible through touch, but not yet that a creature capable of perceiving solidity *must* be equipped with a sense of touch. Someone sceptical of this claim might hold that a subject without a sense of touch could perceive solidity, possibly through visual perception<sup>24</sup>. If this represents a possibility, then the solidity argument is in trouble. Such a possibility is denied by Joske, 1967. In what is, effectively, his own version of the solidity argument, he claims that,

A purely contemplative being could observe visual objects behaving like solid objects behave. Thus, he could observe them moving, falling apart or changing their shape when coming into contact with other objects, and he could also observe that moving objects stopped or veered when they encountered immobile objects. He might describe all objects which behaved in this way as, say, 'apenetrable'. But this concept of apenetrability would not be the same as our concept of impenetrability or solidity (Joske 1967, p.19)

Joske takes solidity to be impenetrability but we can read our own version of solidity into his claim. His claim might be rewritten as being that visual perception could, of itself, only provide an experience as of shmechanical forces, but not mechanical forces. But the passage is less an argument than a statement of the conclusion. The question which immediately arises is, why could this purely contemplative being not possess the concept of mechanical force? The answer given by Joske is that he would not be able to tell the difference between impenetrable and apenetrable objects. He claims that such a subject, "would never be in a position to know that any particular object was itself a solid" (Joske, 1967, p.20). But this point is not enough to show that he could not possess the concept of mechanical force. Being unable to tell the difference

---

<sup>24</sup> Bennett, 1971, p.93, interprets Hume as claiming that solidity can be both tactually and visually perceived.

between *F*s and *G*s does not entail that one does not possess the concept of a *G*. Not being much of a zoologist I cannot perceptually distinguish between ferrets and weasels, but this does not prevent me from possessing the relevant concepts<sup>25</sup>. Furthermore, it seems reasonable to think that I can have an experience as of a ferret. If this is correct, we have no reason to think that the purely contemplative being could not have experience of mechanical force, and thus experience as of physical objects.

Pace Joske, it is highly intuitive to think that one can perceive mechanical force visually, as when one sees a car crash, or a game of snooker. And this point tells against the solidity argument. For the claim that the perception of solidity requires the sense of touch is simply offered with little in the way of justification. Since this is the case, it is of no help to the proponent of the solidity argument that the sense of touch requires bodily-awareness.

Cassam, 2002, offers two replies to the current objection to the solidity argument, the first of which comes in two parts. First, he claims that the fact that a subject can visually perceive mechanical force is not enough to sink the argument, since visual perception itself requires bodily-awareness. He claims that, “Visual experience involves visual sensations, and visual sensations are, like all sensations, bodily occurrences” (Cassam, 2002, p.325). Furthermore, visual experience is integrated with the awareness that we have of the position of the head and its relation to the rest of the body. But this, as far as I can tell, is not to the point. The fact that visual sensations (if such things there be) are bodily occurrences is neither here nor there. Visual sensation does not thereby become a form of bodily-awareness. Furthermore, it is uncontroversial that *our* visual experience is integrated with bodily-awareness in a variety of ways. What the proponent of the solidity argument must claim, however, is that visual experience is *necessarily* integrated with bodily-awareness. And we are not offered any reason to believe this<sup>26</sup>.

The second of Cassam’s replies to the purely contemplative being objection is best seen, I suggest, as a way of bolstering the first reply. He writes,

---

<sup>25</sup> This point relies on the rejection of something like Russell’s Principle, thought of as applying to properties. I am denying the claim that to think of something as *F* one must have discriminating knowledge of *Fness*.

<sup>26</sup> Cassam draws on points made by Ayers, 1991, Vol.I, p.187. But Ayers takes himself to be merely describing the character of our actual experience.

only a creature with a sense of touch and the ability to exert force with its own body can properly be said to see other bodies as exerting and being subject to mechanical forces. The suggestion here is that were it not for one's own bodily engagement with the world, there would be nothing in virtue of which it would be correct to say that force enters into the content of visual experience. (Cassam, 2002, p.325)

Here Cassam suggests that the reason why visual perception is necessarily integrated with bodily-awareness is that visual perception is necessarily bound up with the capacity for action, and that this capacity requires bodily-awareness. I leave discussion of this argument to the following chapter, where I evaluate the merits of, what I call, the action argument.

#### **7.4 Objections**

I have already discussed several objections to the solidity argument. First, it was pointed out that there seems nothing incoherent about the possibility that a subject might enjoy spatial perception but not of a single unified spatio-temporal system. Second, it was objected that place reidentification could occur on the basis of feature reidentification. Neither of these objections, it was argued, poses much of a threat to the solidity argument. However, it was then objected that one might think of objects as subject to mechanical force even though one had never had any experience as of mechanical force. And, finally, even if this objection could be met, it has been suggested that the perception of mechanical force might be available purely visually. Unless it rests on the, hitherto unexplored, action argument, the claim that perceiving mechanical force is only possible for a subject equipped with touch was found to be unjustified. Before I leave the solidity argument, I want to mention two more objections that will, no doubt, have been in the minds of some readers throughout the previous sections.

The first objection is an extremely natural one, and can be raised against any argument which purports to establish that bodily-awareness is a transcendental condition of self-consciousness (or pretty much anything else). The person who makes this objection simply points out that there are living breathing and, most importantly, self-conscious people who lack bodily-awareness. There are people unfortunate enough to have lost the majority of what I have been referring to as

bodily-awareness<sup>27</sup>. These people can, of course, still see their bodies, but they can no longer feel them from the inside. The most striking thing about such people is the amount that they *can* do rather than what they cannot. In particular, they can think of and experience objects as solid, reidentify places, think of themselves as spatially located, and think first-person thoughts. Does the existence of actual cases of body-blindness not show that any argument which attempts to show that bodily-awareness is a transcendental condition of self-consciousness must fail?

The answer to this question must be no. The actual cases of people who have lost their faculty of bodily-awareness have done precisely that, *lost* it. Throughout the discussion of the solidity argument we should be careful not to stretch the conclusion beyond the claim that a self-conscious subject must have had experience of embodiment *at some point*. It is consistent with this claim that a subject might lose their capacity for bodily-awareness and yet remain self-conscious, that mode of thought having already been made available to them. Thus, these actual cases do not show that the solidity argument is bound to fail<sup>28</sup>.

Let us suppose that all of the objections considered thus far can be answered. There remains one last serious objection. The final worry concerning the solidity argument involves the charge of equivocation. Recall that the first step in the argument involved the claim that the reidentification of places involves the reidentification of things, or physical objects. This claim was defended through a defence of the claim that the reidentification of places necessarily rests on the reidentification of things that are thought of, or experienced, as internally causally connected. But, as I pointed out in chapter four, although it may well be that thinking of, or experiencing, something as internally causally connected is a condition of thinking of, or experiencing, something as a physical object, it does not appear to be a sufficient condition. For one may be thinking of, or experiencing, something as internally causally connected, and yet thinking of, or experiencing, it as a Cartesian soul, for instance. Thus, the most we should conclude from this defence of the first step of the solidity argument is that the reidentification of places necessarily rests on the reidentification of *things*.

---

<sup>27</sup> See Sacks, 1985, Ch. 3; Cole, 1991; and Cole and Paillard, 1995.

<sup>28</sup> Bermúdez, 1995, gives an interesting account of the sort of interplay one can expect between actual cases and transcendental arguments of this type.

But, now, when we turn to the next step of the argument, we can see a much richer conception of physical object in play. For here, the argument rests on the claim that being subject to mechanical force is a necessary property of physical objects. But unless this is to count as an equivocation over two notions of object, we would need a linking premise to the effect that thinking of, or experiencing, something as internally causally connected means thinking of, or experiencing, it as a physical object *in this strong sense*. But this premise is not defensible. For it seems as though the reidentification of places need not employ the notion of mechanical force at all. As long as the subject thinks of, or experiences, the reidentifiable particulars as having their later states dependent on their earlier states, nothing more is required. And this appears to be possible. Consider a subject orienting himself by observing a series of holograms which he explicitly takes to be subject only to schmechanical force. There appears to be nothing preventing that subject from thinking of, or experiencing, those as things, i.e. as internally causally connected. But he is by hypothesis not thinking of, or experiencing, them as physical objects, i.e. as being subject to mechanical force.

Given this long list of objections to the solidity argument, it would seem wise to move on. For we are not really able to make any empirical assumptions with which to bolster the argument, without rendering the conclusion trivial. The first empirical assumption which we made at the beginning was that we are only to consider those self-conscious subjects who have experience as of a single unified spatio-temporal system. This seemed innocent enough. But, if we were to patch up every hole in the argument, we would need to assume that the self-conscious subjects that are covered by this argument have experience as of solid objects, and have *tactile* experience as of solidity. But this is to assume so much as to make the conclusion trivial.

## 8. Action

*Q: Why didn't the ghost go to the Halloween Ball?*

*A: Because he had nobody to go with.*

There are two versions of the action argument; an experience-directed and a belief-directed version. Both claim that there is a tight connection between spatial perception and the capacity for spatial action. The experience-directed version then goes on to claim that experience as of having a body is a transcendental condition of the capacity for spatial action. The belief-directed version, on the other hand, claims that thinking of oneself as having a body is a transcendental condition of spatial action. I shall begin by examining the claim that spatial perception requires the capacity for spatial action.

### 8.1 Perception and Action

Several philosophers have claimed that there is a conceptual connection between spatial perception and action. The central claim tends to be,

**(Act)** Spatial perception would not be available to a subject in the absence of any capacity for (spatial or bodily) action.

Philosophers who have endorsed claims very much like **Act** include Schopenhauer, 1966, Merleau-Ponty, 1962, Taylor, 1978-9<sup>1</sup>, Hampshire, 1959, Evans, 1982, Brewer, 1992, and Campbell, 1993. **Act** is an intriguing claim and has some intuitive plausibility, yet this is marred by the obscurity of the arguments which have been offered in its defence. My aim is to offer a defence of **Act** which, so far as is possible, avoids such unclarity. The topic is deep, and the waters murky. But I hope to move some way in the direction of making **Act** look as plausible as possible<sup>2</sup>.

---

<sup>1</sup> Taylor never makes it clear whether he endorses **Act** or whether he is simply offering an interpretation of Merleau-Ponty. I shall speak as though the former were true, referring to it as 'Taylor's argument'.

<sup>2</sup> One philosopher who explicitly rejects **Act** is G. Strawson, 2003, p.228.

In considering the views of both Schopenhauer and Merleau-Ponty, special interpretational difficulties arise. Since this is the case I take, as my point of departure, an argument offered by Taylor. This seems like a sensible place to begin, since a direct line of influence can be traced from Taylor to Evans, and from Evans to both Brewer and Campbell.

Taylor's argument for **Act** can be summarised as follows: To begin with, "our perceptual field has an orientational structure, a foreground and a background, an up and down. And it must have; that is, it can't lose this structure without ceasing to be a perceptual field in the full sense" (Taylor, 1978-9, p.154). The particular feature of the perceptual field that Taylor concentrates on is the up/down dimension<sup>3</sup>. According to Taylor, the perceptual field has the structure it has (i.e. has an up/down dimension) "because it is experienced as a field of potential action" (p.155). That is, the perceptual field has its up/down dimension because the world is experienced through perception, as a place in which we can act. Why is this? The answer, says Taylor, is that "up and down are related to how one would move and act in the field. For it is only as a bodily agent functioning in a gravitational field that 'up' and 'down' have meaning for me." (pp.154-5)<sup>4</sup>. Thus, we come to the conclusion that "our perception of the world is essentially that of an embodied agent, engaged with, or at grips with the world." (p.154).

We can go through this argument step by step. First we must ask whether it is plausible to say that the perceptual field necessarily has an up/down dimension. It might seem so. If there were no experienced up or down, then the field of perception would lack one of its dimensions, and that does not appear to make sense. We simply can't imagine what it would be like for our perceptual field to lack the up/down dimension. The absence of the up/down dimension, as it were, shrinks out the other dimensions. Think of the image of the television screen thinning to a horizontal line before finally vanishing. If this is right, then we might be persuaded to agree with Taylor, and say that if our perceptual field lost its up/down dimension, it would indeed cease 'to be a perceptual field in the full sense'.

---

<sup>3</sup> Use of the phrase 'perceptual field' should not be thought to imply a sense-data theory of perception. See Martin, 1992.

<sup>4</sup> Compare, "there is no way of specifying what 'left' say, means, except by pointing to the left." (Eilan, 1997, p.237).

But we have moved too quickly here. Certainly it is difficult, perhaps impossible, to imagine visual perception without the up/down dimension, but what of hearing? Hearing is a form of spatial perception, but we do not seem to have much difficulty in imagining ourselves lacking the up/down dimension in auditory perception. It will immediately be objected that hearing only gives us spatial information because it is so intimately related, in our own cases, with tactile, kinaesthetic and visual perception<sup>5</sup>. It would follow that the possibility of imagining auditory perception lacking the up/down dimension is of little importance, since it only qualifies as spatial perception in some derivative sense. This may well be true, but there is, in any case, a stronger objection to Taylor's claim that if our perceptual field lost its up/down dimension, it would cease 'to be a perceptual field in the full sense'.

The objection is that Taylor's view is anthropomorphic. Consider the case of Eyeball. Eyeball is an intelligent creature inhabiting some, not so remote, possible world. Eyeball is rather like a slug, with the major difference being that his body consists in a giant, perfectly spherical eyeball, through which he can see in every direction at once. Eyeball is able to perceive his environment, and move around in it accordingly. Eyeball has the advantage that no-one can ever sneak up behind him, but the disadvantage that he can never turn his back on anybody. I have no hesitation in saying that none of us can know, or even imagine, what it is like to be Eyeball. We simply have no way of imagining what Eyeball's visual experiences are like. But we should not doubt that Eyeball *has* visual experiences. The question is whether Eyeball's visual field is structured with an up/down dimension. I see absolutely no reason to think so<sup>6</sup>. The point is that even though we cannot imagine what visual experience would be like without the up/down dimension, we have reason to think that it is at least possible. And if this is right, Taylor's claim cannot be accepted<sup>7</sup>.

---

<sup>5</sup> As is claimed by Strawson, 1959, p.65.

<sup>6</sup> Won't Eyeball still feel the pull of gravity and so orientate his world into an up and down? No – Eyeball lives in space. But doesn't Taylor admit that, "it is only as a bodily agent *functioning in a gravitational field* that 'up' and 'down' have meaning for me" (Taylor, 1978-9, p.154-5, emphasis added). Well, if he *really* thinks that, he must think that perceptual experience is only possible for creatures that have grown up with gravity. This is implausible.

<sup>7</sup> It might be suggested that Taylor's argument is not drastically undercut by this objection. For it may be that Taylor is interested in the necessary conditions of having spatial perception *of the kind we have*. For we have allowed that any spatial perception imaginable by us has an up/down dimension. This reading is supported by Taylor's claim that "we couldn't have a subject with a field *articulated like ours* who as a matter of contingent fact, might not be an embodied agent" (Taylor, 1978-9, p.156, emphasis added).

This objection, which I believe to be a good one, is an application of a more general point made by Campbell. Campbell's aim is to define the notion of an egocentric frame of reference. In this connection, he writes that "an approach which tries to define what it is for a frame to be egocentric by simply listing a set of axes will not work. Again, there is no reason to suppose that all species will use the same egocentric axes" (Campbell, 1993, p.74). Campbell's alternative to this 'extensional approach' to defining the notion of an egocentric frame of reference is to consider "how the subject is apprehending those axes" (p.74). That is, what is distinctive of an egocentric frame of reference is not the axes used, but the way in which those axes are apprehended by the subject.

We might make a similar point as regards Taylor's argument. Instead of claiming that the perceptual field necessarily has a certain orientational structure (i.e. an up/down dimension), we should consider the possibility that the orientational structure of the perceptual field, whatever that may be, is necessarily apprehended in a particular way. We might, then, try to salvage the bones of Taylor's argument by claiming that: (1) the perceptual field necessarily has an orientational structure, (2) this orientational structure is necessarily apprehended by the subject as a 'field of potential action', for (3) it is only 'as a bodily agent functioning in a gravitational field' that the structure of the perceptual field has 'meaning for the subject'. Therefore, (**Act**) spatial perception would not be available to a subject in the absence of any capacity for (spatial or bodily) action.

All this is still rather obscure, but if it can be made satisfactorily transparent, at least we will have managed to evade the charge of anthropomorphism. Eventually, I want to endorse something very much like this defence of **Act**. Clearly this can't be done until we have some elaboration on the enigmatic (3). Before I get on to (3), however, I want to say something in explanation and defence of (1).

I take (1) to mean, roughly speaking, that perception is necessarily from a particular point of view. Perception is perception *from somewhere*. Furthermore, the content of perception is, what we can call, self-locating. That is, in perceiving the world I am simultaneously made aware of where worldly objects are and where I myself am. As Brewer puts it, "perception places us in the perceived world" (Brewer, 1992, p.17). This seems right. If to perceive something is to perceive it from

---

somewhere, some of the content of any perceptual state will represent the spatial relations which hold between the perceiver and the objects perceived; it is 'egocentric'. If perception is always perception from somewhere, it seems inconceivable that the content of a perceptual state could fail to include some information as to the position of the perceiver. So, we should agree with Taylor that the perceptual field necessarily has an orientational structure<sup>8</sup>.

What remains to be seen is whether any sense can be given to the claim that it is only 'as a bodily agent functioning in a gravitational field' that the structure of the perceptual field has 'meaning for the subject'. In giving a sense to this claim we will, I suggest, be answering the question, 'In virtue of what is perception self-locating?' or, in Taylor's terms, 'In virtue of what does the perceptual field have its orientational structure?'. We must approach these questions carefully, through a proper understanding of the notion of a frame of reference. For, yet another way to ask what is the question is, 'What is it for a frame of reference to be egocentric?'. To answer these questions, we must look at the notion of a frame of reference. In the following two sections we will see that Campbell is right to say that listing a set of axes will not help us here. What is more, we shall see that listing particular *types* of frames of reference will not help us either. Rather, we need to look at how a particular set of axes, or type of frame of reference, is apprehended by the subject.

## 8.2 Frames of Reference

The idea that spatial representations (be they linguistic, pictorial, perceptual or whatever) employ Frames of reference (*FoRs*) is a common one. The notion of a *FoR* is best approached by example. Consider the following perceptually presented scene showing a man and an ambulance:

---

<sup>8</sup> There is a question as to whether perceptual content is self-referential or self-indicating, i.e. is the content of perception 'door ahead of me' or 'door ahead'? On the self-referential view, perceptual content contains the *I* concept. On the self-indicating view, perceptual content does not contain the *I* concept, but has content which justifies the corresponding *I* judgement. This is thought to be so because the perceptually presented 'door ahead' a priori entails 'door ahead of me' for those with the relevant concepts. For the self-referential view, see Peacocke, 1999, p.264. For the self-indicating view, see Campbell, 1994a, p.119.



fig.3

Here we can say that the man is to the left of the ambulance. However, we can also say that the man is in front of the ambulance, without implying that either has moved. The reason is that in each case we are employing a different *FoR*. We might think of a *FoR* as a way of specifying the spatial relations between objects. More specifically, a *FoR* is an origin and set of axes which define a direction system. Thus, relative to one origin and set of axes, the man is to the left of the ambulance, relative to another he is in front of it. When we say that the man is in front of the ambulance, it is natural to think that the axes employed are the *left/right, front/back* axes, and the origin is the ambulance. When we say that the man is to the left of the ambulance, it is natural to think that the axes employed are the *left/right, front/back* axes, and the origin is the viewer<sup>9</sup>.

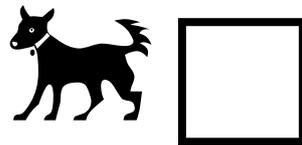
With these notions introduced, we can move on to a more elaborate characterisation of the different kinds of *FoR*. Unfortunately there are several competing views on how to distinguish between different kinds of *FoR*. Instead of reviewing the literature on frames of reference, I will simply outline the view that seems, to me, to make most sense of the debate. This is the framework proposed by Levinson, 1996. This framework is described at the level of linguistic representation, so we must briefly turn to language. We'll return to perception later.

Levinson distinguishes between three kinds of *FoR*: intrinsic, absolute and relative. Let me work through each in turn. An intrinsic *FoR* is an object centred set of axes, with co-ordinates determined by the intrinsic features of the object. A spatial relation described by an intrinsic *FoR* can be called an intrinsic spatial relation. Intrinsic spatial relations are two-place relations, holding between the object upon which the axes are centred (ground) and the object who's position is determined relative to it (figure). For instance, in the above example, the sentence 'the man is in front of the ambulance' employs an intrinsic *FoR*. The set of axes are centred on the

---

<sup>9</sup> According to the framework proposed below, these intuitive analyses turn out to be slightly misleading.

ambulance (ground), and co-ordinates are determined by its intrinsic features, i.e. its front and back, and its left and right-hand sides. Any object which has intrinsic features suitable for defining such co-ordinates can act as the origin of an intrinsic *FoR*. For instance, concerning *fig.4*, ‘the cube is behind the dog’ employs an intrinsic *FoR* whose origin is the dog. This is because the dog has an intrinsic front, back, left and right. However, ‘the dog is behind the cube’ does not employ an intrinsic *FoR* whose origin is the cube. This is for the simple reason that the cube doesn’t have the intrinsic features required to specify its back<sup>10</sup>. If it could ever be true to say that the dog is behind the cube, there must be some other (non-intrinsic) *FoR* in play.



*fig.4*

Next we can turn to absolute *FoRs*. There is a sense in which spatial representations which employ an absolute *FoR* are more simple than those which employ an intrinsic *FoR* since we don’t have to concern ourselves with the intrinsic features of objects. An absolute *FoR* makes use of some fixed set of co-ordinates. An example is the *up/down* axes determined by the force of gravity, another is the set of axes comprising *north/south/east/west*. Absolute *FoRs* specify the spatial relation that holds between two objects (figure and ground) in terms of these ‘external’ fixed co-ordinates. For instance, if we consider *fig.5*, we might say that ‘the weightlifter is to the east of the car’. Here we are employing an absolute *FoR*. The axes used are the *north/south/east/west* axes, and the origin is the car. The sentence ‘the car is to the west of the weightlifter’ employs the same set of axes, but with its origin on the weightlifter.

---

<sup>10</sup> Whether a given object is regarded as having such intrinsic features varies across different cultures. This is made clear in Levinson, 1996. For instance, the English language does not recognize an intrinsic front or back for objects such as trees. This differs from some Nilotic cultures, who say that the front of a tree is the opposite of whichever way it leans. See Levinson, 1996, n. 43. To this extent it might be regarded as infelicitous to call such features intrinsic, since we do not usually think of an object’s intrinsic features as relative in this way. In what follows I shall ignore this worry.



fig.5

A spatial relation described by an absolute *FoR* can be called an absolute spatial relation. Absolute spatial relations are two-place relations holding between figure and ground. As with intrinsic *FoRs*, the origin of the axes is always centred on the object which is the ground.

Spatial representations which employ relative *FoRs* display the most complexity. Relative *FoRs* differ from both the intrinsic and absolute in that the spatial relations they describe are three-place relations. They involve, not only a figure and ground, but also a viewpoint. They use axes centred on the viewpoint to assign positions to both figure and ground. For example, the sentence ‘the man is to the left of the ambulance’ is true of *fig.3*. But this clearly is not employing an absolute *FoR*, neither is it employing an intrinsic *FoR*, since the man is to the intrinsic front of the ambulance. Rather, it employs a relative *FoR*. It is true that the man is to the right of the ambulance, but only *relative to the viewpoint of the speaker*.

The viewpoint which is the origin of the axes of a relative *FoR* need not be that of the speaker. If we look at *fig.6*, and imagine A uttering ‘from B’s point of view, the cat is to the right of the fish’, he is employing a relative *FoR* who’s axes are centred on the viewpoint of an individual other than himself.

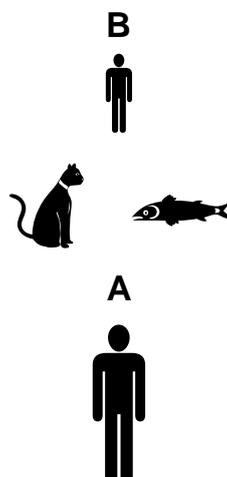


fig.6

Levinson's suggestion is that every linguistic spatial representation can be analysed in terms of these three kinds of *FoR*. All that is required is grasp of the notions of a set of axes, an origin, a figure, a ground and a viewpoint. It seems to me that Levinson's framework is both clear and useful. It will become obvious, however, as we return to discuss *FoRs* in perception, that there are other distinctions that must be brought into play.

Assuming that perceptual states represent the world as being a certain way spatially, and that these representations employ *FoRs*, we can ask what kind of *FoRs* are used<sup>11</sup>. Is the full range of kinds of *FoR* employed in perceptual spatial representation? Let us go through each in turn. It seems reasonably plausible to suggest that perception can employ intrinsic *FoRs*. In *fig.3* we see that the man is standing in front of the ambulance. Surely 'the man is in front of the ambulance' is part of the representational content of the perceptual state. If this is the case, perception employs intrinsic *FoRs*.

What about absolute *FoRs*? It seems pretty clear that we can see that an object is above another. This representational content uses an absolute *FoR*. 'Above' is rarely used in any *FoR* other than an absolute one with axes determined by the force of gravity. 'Above' can be used to mean 'above from my point of view' (relative), or 'to the intrinsic top' (intrinsic), but such uses are rare. For instance, it is very natural to say that, no matter what angle it is viewed from, the bird in *fig.7* is above the radio. Even if the radio were upside down, we would still say that the bird was above it.



*fig.7*

---

<sup>11</sup> The points I make about frames of reference in this chapter are made in terms of a representational theory of perception. I assume, however, that the proponent of a non-representational theory can translate this talk into his or her favourite idiom.

Gravity is perceptually detectable. We feel its pull, and see it in action. This gives credence to the claim that perceptual content can employ absolute *FoRs*. But it might seem that there is no perceptual representation of absolute spatial relations on the horizontal plane. In English it is the *north/south/east/west* directional system that is most widely used as the external, fixed set of co-ordinates. And we don't *see* that some object is, for instance, to the west of another. But this view becomes less obviously correct when we consider the particular characteristics of some other languages. Levinson points out that Tzeltal (the relevant dialect of which is spoken by about 15,000 people in the Tenejapa community in Chiapas, Mexico) employs an *uphill/downhill/across* system. The community lives on the side of a mountain range, and these fixed axes, which are used in absolute *FoRs*, are readily perceptually detectable. Does it not seem plausible to say that a native Tzeltal speaker can see that a tree, say, is 'downhill' of a house? I see no obvious reason for denying that the representational content of such a person's perceptual state might be expressible by 'the tree is downhill of the house'. So, it seems likely that spatial perception can and does employ absolute *FoRs*.

It seems obvious enough that perception employs relative *FoRs*. It is the most natural thing in the world to say that we see *a* to the left of *b*, or in front of *b*, where this is not employing an intrinsic *FoR*. Looking at *fig.3*, it is part of the representational content of my perceptual state that the ambulance is to the right of the man. Furthermore, one might even argue that, concerning *fig.6*, the following can be represented in A's perceptual state, 'from B's point of view, the cat is to the right of the fish'. Of course it will be controversial whether perceptual representations can have such rich content (involving points of view etc.).

The conclusion seems to be that the perceptual representation of spatial relations employs intrinsic, absolute and relative *FoRs*. But things are not quite so simple. Consider the perceptual state I am in when I stand, facing forward with my eyes open, looking at a tree. In reporting the content of my state, I can say either 'there is a tree in front of me' or 'I am in front of a tree'. The first of these sentences employs an intrinsic *FoR*, with the set of axes centred on me. But the second does not. The *FoR* is not intrinsic, since the tree has no intrinsic front. The *FoR* must be relative, with myself as both viewpoint and figure. That is, I am in front of a tree, relative to my own point of view.

But it is unclear that this distinction should, or even can, be made when it comes to specifying the content of the perceptual state. It would seem to be an obvious case of double counting to say that my perceptual state has, as part of its content, *both* that I am in front of a tree, and that there is a tree in front of me. Thus, a note of caution is raised. Perhaps perceptual representation is at a more basic level than linguistic representation, in the sense that the intrinsic/absolute/relative distinctions don't easily apply. Of course, some philosophers will say that this phenomenon is simply a consequence of the fact that perceptual content is non-conceptual, and can be conceptualised in a number of equivalent ways. And, of course, this may very well be the right answer. I don't think that this view is forced upon us, however, for it is perfectly open for the conceptualist to claim that perceptual content is conceptual, but at a level which doesn't always respect the distinction between relative and intrinsic ways of representing spatial relations.

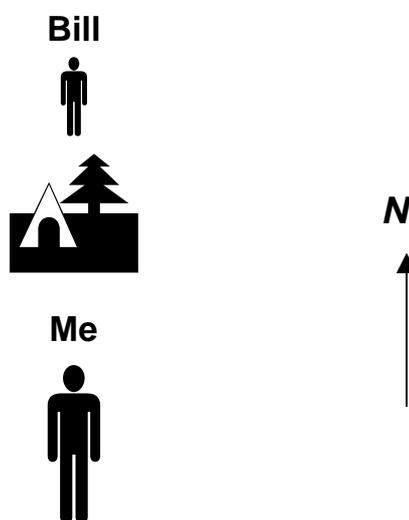
If perceptual representations employ all three kinds of *FoR*, are we any closer to saying in virtue of what it is that perception self-locating? It is intuitive to think that we are. For it is tempting to think that the perceptual representation of space is special in that it employs *FoRs* that are relative. Perception provides one with information about the spatial location of objects around one *relative to one's own point of view*. Unfortunately, we cannot be satisfied with this answer. To see this consider a map, the paradigm of a spatial representation which is not self-locating. Think of this map as a photograph taken by a satellite positioned directly above you. The photograph contains you as an element, as it does the various things around you including, say, other people. This map represents the spatial relations between different objects in your environment and also between those objects and yourself. Furthermore, it conveys information in intrinsic, absolute and relative *FoRs*. Yet this map is not self-locating. It does not, of itself, tell you where you are. Pointing out that perception uses *FoRs* which are relative (as well as intrinsic and absolute) does not tell us what it is in virtue of which perception is self-locating.

### **8.3 Viewer-Centred Frames of Reference**

The phrase 'egocentric frame of reference' is often used to mean 'frame of reference centred on the viewer/speaker'. I want to stay away from this use of the term 'egocentric'. Hence I shall use the phrase 'viewer-centred frame of reference'. It is

sometimes claimed that perceptual states necessarily represent spatial relations in a viewer-centred *FoR*<sup>12</sup>. A viewer-centred *FoR* is any *FoR* who's origin is, or is a part of<sup>13</sup>, the speaker/thinker/perceiver. A non-viewer-centred *FoR* is any *FoR* who's origin is not (any part of) the speaker/thinker/perceiver.

According to this method of classifying *FoRs* any kind of *FoR* (intrinsic, absolute, or relative) can be viewer-centred. Linguistic examples (see *fig.8*) are 'the tree is in front of me', 'the tree is north of me', and 'the tree is to the right of the wigwam' respectively. Similarly, all three kinds of *FoR* can be non-viewer-centred. Linguistic examples are 'the tree is in front of Bill', 'the tree is to the south of Bill', and 'from Bill's point of view, the tree is to the left of the wigwam'.



*fig.8*

Is it correct to claim that perception necessarily employs viewer-centred *FoRs*? There are two ways of understanding this claim. On one understanding I take it to be almost certainly false, on the other it is very close to being an important truth. The first (almost certainly false) of these is to understand the claim as meaning that every perceptual representation of space employs a viewer-centred *FoR*. If this were true, it would entail that we cannot see that the tree is in front of Bill. For that employs a non-viewer-centred *FoR*. Similarly, consider *fig.3* again, it would have the consequence

<sup>12</sup> See, for instance, Bryant, 1997, pp.247-8. Although he uses the term 'egocentric' I think it is clear that he means what I mean by 'viewer centred'.

<sup>13</sup> The general consensus is that in order to give an adequate explanation of the way in which spatial representations feed into action, it is necessary to employ a variety of viewer-centred *FoRs* centred on different parts of the body. See the various papers in Paillard, 1991.

that we cannot see that the man is in front of the ambulance. I take these consequences to be highly counterintuitive and suggest that, on this way of understanding it, the claim that perception necessarily employs viewer-centred *FoRs* is false.

But there is another way of understanding the claim. We could understand it as the claim that every perceptual state must have at least some content which employs a viewer-centred *FoR*. Typically, every perceptual state has a vast amount of representational detail. So, for example, if presented with *fig.3*, we see that the man is in front of the ambulance, and that the ambulance is to the right of the man, and that the ambulance has a cross on it, and so on. On this second way of understanding the claim, although some of the content of a perceptual state may employ non-viewer-centred *FoRs*, not all of it can. There must be some part of the content which employs a set of axes whose origin is the perceiver. This seems plausible. It stems from the fact that perception is necessarily perspectival. To perceive something is to perceive it from somewhere. Thus, some of the content of any perceptual state will represent the spatial relations which hold between the perceiver and the objects perceived. As was pointed out earlier, if perception is always perception from somewhere, it seems inconceivable that the content of a perceptual state could fail to include some information as to the position (point of view) of the perceiver.

True as this may be, it does not help us to understand what is so special about the way in which perception represents spatial relations. To see this we need to consider, what might be called, Perry cases<sup>14</sup>. A Perry case is a case of someone who perceives or refers to himself without realising it. The most famous example is that of Oedipus. If Oedipus were to utter the words ‘Jocasta is standing behind the slayer of Laius’, he would be referring to (or designating) himself but without realising it. Indeed, he would be employing an intrinsic, viewer-centred *FoR*, but without realising it. Turning to the perceptual case, we can imagine someone who doesn’t realise that he is looking into a mirror, seeing that the bogeyman is behind the rather good looking fellow. Here his perceptual state employs an intrinsic, viewer-centred *FoR*, but he doesn’t realise it.

These cases show us that the current proposal, that all perceptual states contain some viewer-centred content, cannot fully account for what is special about the

---

<sup>14</sup> See Perry, 1979.

perceptual representation of spatial relations. For it is entirely consistent with this proposal that all the viewer-centred content of a perceptual state be like a Perry case. But surely this is inconceivable. Unlike the map we considered in the last section, perception just isn't like this. The idea that a subject might have a perceptual awareness of the world and yet be unable, on the basis of that perceptual awareness, to locate him or herself is not conceivable. It would not be spatial perception in the sense that we are familiar with. Indeed, even if it were possible, it would be insufficient to satisfy the demands of the self-location argument, for that argument required that the subject be able to locate him or herself in space. So, in order to capture what is special about the self-locating perceptual representation of space, we are going to have to say more than simply that it necessarily employs viewer-centred *FoRs*.

#### 8.4 Egocentric Frames of Reference

For an *FoR* to be egocentric doesn't seem to be a matter of the kind of axes used, or a matter of where they are centred. Thus, Campbell's claim, that the 'extensional' approach is misguided, is vindicated. But what is it for an *FoR* to be egocentric? The answer, according to the action argument, lies in the intimate connection between the content of spatial perception and action. Egocentric *FoRs* are those which contain information that can be used immediately in action. Consider a fact, noted by Evans<sup>15</sup>, that, "egocentric spatial terms are the terms in which the content of our spatial experiences would be formulated, and those in which our immediate behavioural plans would be expressed" (Evans, 1982, p.154).

Suppose I see a glass on a table in front of me. The part of the content of this perceptual state which specifies the spatial relation between me and the glass<sup>16</sup> is immediately usable by me in action, i.e. picking up the glass. For an intention to pick up the glass will involve exactly that content. This clearly seems to be the case. The next move is to use this fact in a definition of egocentric *FoRs*. For this fact leads naturally to the suggestion that an egocentric *FoR* is one which provides the subject with information that can be used immediately in the direction of action. As Campbell

---

<sup>15</sup> Also see, Peacocke, 1992, §4.

<sup>16</sup> Or maybe, my hand and the glass.

writes, “the axes that are distinctive of an egocentric frame are those which are immediately used by the subject in the direction of action.” (Campbell, 1993, p.75).

This marks a clear difference between the way in which perception represents the perceived environment, and the way in which a map does. The spatial relations between myself and my environment represented in the satellite photograph map mentioned earlier, cannot be used immediately in the direction of action. For, I need also to know *which thing on the map is me*. Suppose the map is of Trafalgar square, from above. I cannot, just on the basis of the information given (without knowing which of the things represented is me), move towards Nelson’s column. However, this is not the case with perception. Nothing more is required than that I see which direction Nelson’s column is in. This, of course, is because perception is self-locating. And, the suggestion is, what it is in virtue of which perception is self-locating, is that it can be used immediately in the direction of action<sup>17</sup>.

We want this to be a necessary and sufficient condition. It seems reasonable to say that for an *FoR* to be egocentric, nothing more is required than that it is immediately usable in the direction of action. For if it is *immediately* so usable, then it really does seem to have ‘placed us in the perceived world’. Those sceptical of the account will argue that the condition is not necessary; that a *FoR* can be egocentric without it being immediately usable in the direction of action. To put some flesh on this objection it would be necessary to give a plausible example of a perceptual state, employing an egocentric *FoR*, but which is not immediately usable in the direction of action. For the time being I am content to conjecture that no such example will be forthcoming. Of course part of the reason for this is that we lack a plausible alternative account of what it is for an *FoR* to be egocentric. But, speaking intuitively, it seems that any representational state which succeeds in ‘placing the subject in the world’ will certainly be *usable* in the direction of action for, by definition, information is given relating the subject to his surroundings, and it is that information which guides action.

I have argued that there is a link between the perceptual field’s having an oriented structure and the capacity for spatial action. But this is not yet complete as a defence of **Act**. For **Act** claims that spatial perception would not be available to a

---

<sup>17</sup> Eilan, 1995, p.348, makes the connection between the claim that perception is egocentric, in the present sense, and Gibson’s claim that we perceive affordances for action.

subject in the absence of any capacity for (spatial or bodily) action. All we have said so far is that spatial perception necessarily provides the subject with information which is usable in the direction of action. But the fact that some perceptual information is usable in the direction of action, does not entail that the subject of such perceptual states is capable of action. We might think that there could be a subject who has egocentric perception, which would be usable by it in the direction of action *were it capable of such action*, but which is itself incapable. And presumably this is the position of someone whose entire body is paralysed.

But, generally speaking, a subject with entire body paralysis was not always like that. Quite obviously, such a subject's capacity for spatial perception can outlast his or her capacity for bodily action. Of course some story will have to be told, but we might think it plausible that such a case is compatible with a liberal reading of **Act**. A better test case is a subject who is *constitutively* incapable of spatial action. Even better, in my opinion, is a subject who has no grasp of the possibility of acting in the space which it perceives.

Consider four subjects: **A**, **B**, **C**, and **D**. **A** is fully capable of spatial action; **B** used to be capable of such action but is now entirely paralysed; **C** has never been capable of spatial action but has some grasp of the notion of spatial action and that it is not a possibility for him; **D** is constitutively incapable of spatial action, and has no grasp of the possibility of its spatially acting<sup>18</sup>.

According to no reading of **Act** is there any reason for thinking that **A** could not enjoy spatial perception. Similarly, it would be an implausibly strict reading of **Act** that entailed that **B** was not capable of spatial perception. **C** and **D** are much more difficult cases. One worry is that it is not immediately clear that **C** and **D** really do represent alternative possibilities. It is not clear that a subject could have a full grasp of the possibility of spatial action if it were constitutively incapable of such action. Let us ignore this worry. More pressing, I think, is that it is very difficult to see upon what grounds we could argue that **C** could not enjoy spatial perception. For even if we agree that there is a link between spatial perception and the capacity for spatial action, such that (i) spatial perception is necessarily egocentric, and (ii) egocentricity is a matter of being immediately usable in the direction of action, there is still a plausible

---

<sup>18</sup> De Gaynesford, 2002, argues against **Act**. However, he fails to adequately distinguish between **B**, **C**, and **D**. His arguments have no force against the minimal version of **Act** endorsed below.

reading upon which **C** is perfectly able to enjoy spatial perception. For having allowed **B** spatial perception (as we must) it seems a small step to allow it to **C**. It seems to make perfect sense to say that **C** grasps that his perceptual information *would be* immediately usable in the direction of action, if only he could move.

It seems to me, then, that we need a minimal reading of **Act**, such that spatial perception would not be available to a subject in the absence of any grasp of the possibility of its spatially acting. Anything stronger looks implausible. But now we owe some answers to some difficult questions: What does it mean to say that a creature does, or does not, grasp the possibility of its acting? Further, why think that (i) and (ii) lead to this minimal version of **Act**?

The answers to these questions come as a package. I shall argue that egocentric representations cannot be attributed to a subject for whom they do not ‘have a meaning’, and that an egocentric content’s having a meaning for a subject is a matter of his grasping the possibility of his acting on that content. Thus, we can lend our support to this minimal version of **Act**. The language used here refers back to that of Taylor. For, as I reconstructed it, Taylor’s argument claimed that it is only ‘as a bodily agent functioning in a gravitational field’ that the structure of the perceptual field has ‘meaning for the subject’. It remains to explain what these things mean.

## **8.5 Having a Meaning for the Subject**

Let’s summarise where we are so far: Spatial perception necessarily gives us egocentric spatial content (employs egocentric *FoRs*). Egocentric content is spatial content that is immediately usable in action. Taylor claims that it is only as an agent that this kind of content has a meaning for the subject. The implication being that in order to ascribe the content to a subject, it must have a meaning for that subject. But what does this mean? And why should we think it true? In an attempt to answer these questions, I shall return to his example of the up/down dimension.

Taylor’s use of the notion of meaning in his argument suggests that he considers it a transcendental condition on my having a perceptual field with the up/down dimension, that ‘up’ and ‘down’ ‘have a meaning for me’. But surely this does not mean that in order to have a perceptual field, I must understand the words ‘up’ and ‘down’. For this would be to deny spatial perception to children and animals. And this just isn’t an option. So what can we interpret Taylor as claiming? I suggest

that we take Taylor as suggesting that the space of the visual field is necessarily intelligible to me. This means that the up/down dimension is necessarily intelligible to me. Taylor's claim can then be reformulated as the claim that the up/down dimension is only intelligible to me given my status as a bodily agent.

But what does this mean? What is it for the up/down dimension to be intelligible to me? Or, more generally, what does it mean to say that the space of the visual field is necessarily intelligible to me? One thing this can't mean is that I must be in possession of the concepts with which the space of the visual field would be specified (i.e. 'up', 'down' etc). For this would, again, be to deny spatial perception to animals, and maybe children. The relevant sense of 'intelligible' is as follows: to find the space of the visual field intelligible is to be able to respond *appropriately* to differences in the content of spatial perception. Responding appropriately is not just a matter of registering different spatial contents as different<sup>19</sup>. Responding appropriately means registering different spatial contents as different *in the very way in which they are different*. This means that being able to appropriately register differences on the up/down dimension requires some grasp, however primitive, of the up/down dimension. This grasp is, without a doubt, not a matter of possessing certain concepts. It is the sort of grasp that lower animals have of the up/down axis. Such animals find the space of their perceptual field intelligible. That is, they are able to respond to differences on the up/down dimension *as* differences on the up/down dimension. This means that they have some rudimentary grasp of up and down. And since such egocentric content is content that is immediately usable in the direction of action, this means having a grasp of the possibility of acting on that content.

So, when Taylor says that 'it is only as a bodily agent functioning in a gravitational field that 'up' and 'down' have meaning for me', I interpret him to mean that it is only as a (bodily) agent that I have a rudimentary grasp of the up/down dimension. Such a grasp is required if I am to find the space of my perceptual field intelligible. 'Up' and 'down' (or any other egocentric terms) are directions of possible action. So, finding my perceptual field intelligible is a matter of experiencing it as a field of potential action. And finding the space of my perceptual field intelligible is indispensable to having such a field at all.

---

<sup>19</sup> See Evans, 1982, pp.154-5.

But what reason do we have for thinking that this final claim is true? Why must spatial perception be intelligible? The reason is that the egocentric nature of spatial perception means that we hear a sound as coming from such-and-such a direction, or we see an object in such-and-such a direction. That the object is in such-and-such a direction is part of the content of those perceptual states. The thought is that if I were unable to respond appropriately to such contents, I could not be ascribed those representations. A creature might be able to discriminate differences in visual information on the up/down axis, but if it were not disposed to look up when it registered information from an object coming from above, what sense can we make of the supposition that it really did see the object *as above it*? It seems that perceiving spatial relations requires being disposed to respond appropriately to differences in content. Such appropriate response requires a primitive grasp of the space as egocentrically presented, and this means experiencing it as a possible field of action. This is what it means to say that spatial perception is intelligible.

The defence of **Act** that emerges from these considerations is as follows: Spatial perception necessarily has an orientational structure. This orientational structure can be thought of as a matter of perception being self-locating, or employing egocentric *FoRs*. An *FoR* is egocentric if it is immediately usable in the direction of (spatial) action. But perceptual space must be intelligible to the subject of perception. This means that a subject must be able to respond *appropriately* to differences in spatial information. This entails that a subject of spatial perception must have some, however primitive, grasp of its egocentric space as a space which is immediately usable in the direction of action. Therefore, spatial perception necessarily gives us content which must be grasped by the subject as immediately usable in action. This is tantamount to saying that the perceptual field is necessarily experienced as a field of potential action. This constitutes my defence of a minimal version of **Act**, the claim that spatial perception would not be available to a subject in the absence of any grasp of the possibility of its spatially acting.

## **8.6 Spatial Action and Bodily-Awareness**

There is both an experience-directed and a belief-directed version of the action argument. In this and the following section I look at each in turn. The experience-directed version of the argument rests on the claim that one can only possess the

capacity for spatial action, or grasp that it is a possibility for one, if one experiences a body as one's own. I shall be arguing against this claim. My view is that bodily-awareness is not even a transcendental condition of the capacity for spatial action, let alone a condition of the ability to grasp the possibility of spatial action. As such, the experience directed argument is a failure.

One philosopher who appears to endorse this step of the action argument is O'Shaughnessy, 1980, and, 2000. He writes,

intentional bodily movings require *awareness* of the animal's body: more, an *attentively immediate* awareness of the primitive variety that obtains in us all [bodily-awareness]. The following makes this clear. Consider a situation in which a man's right arm does not in this primitive way seem to its owner to be postured in any way, indeed does not in this way even seem *there*. Then irrespective of whether he vividly sees his arm, and knows exactly how it is spatially disposed, such a man is not in a position to immediately move it. Although he is not paralysed, in these circumstances his bodily will is simply unavailable for use upon his right arm. Thus he can instrumentally move his right arm with his left arm, he may even be able to instrumentally move it via a mental act like visualising, but he will not be able to immediately move it. This is because the bodily will takes an immediately given bodily object. What else? (O'Shaughnessy, 2000, p.659)

The idea is that intentional action (of the spatial variety) is only possible because the body-part to be moved is *there* for one. In the words of O'Shaughnessy's earlier book, it is *immediately present*.

I suggest, then, that a necessary condition, though not of course a sufficient, of willing a phenomenon  $\phi$  in a limb L, is that in some sense of the expression  $\phi$  and L be 'immediately present' to one. (O'Shaughnessy, 1980, p.145)

If, as in the case of O'Shaughnessy's example, one's arm was not immediately present to one, one would not be in a position to immediately move it; to will a movement in it. The conclusion being that without bodily-awareness, no spatial action would be possible. Since bodily-awareness is, centrally, a way of experiencing a body as one's own, this last step of the action argument is secured relatively easily.

All this seems rather plausible, but notice a complication concerning the distinction between basic and non-basic (instrumental) actions. A basic action is an action *a* that one must perform in order to perform any other action, and for which there is no action that is performed in order to do *a*. An instrumental action is an

action  $a^*$  for which there is a basic action which was performed in order to do  $a^*$ <sup>20</sup>. For example (assuming, as O'Shaughnessy does, that bodily movements are basic actions), moving one's fingers is a basic action, turning the key, and opening the door are both instrumental actions. Basic actions are things that we do *immediately*, non-basic actions are things that we do *non-immediately* or instrumentally.

Now, it should be noted that both of the quotations from O'Shaughnessy, above, concern basic actions. The first quote contrasts moving an arm immediately, with moving it instrumentally. The second claims that bodily awareness is a necessary condition of *willing* a limb to move. And, according to O'Shaughnessy's usage, one can only *will* basic actions. Thus, it is not being claimed that a limb's being immediately present to one is a necessary condition of one's moving it, but that its being immediately present is a necessary condition of one's *immediately* moving it, i.e. one's performing a basic action with it.

Now consider the following passage concerning the man who's arm is not immediately present, "he can instrumentally move his right arm with his left arm, he may even be able to instrumentally move it via a mental act like visualising, but he will not be able to immediately move it". This is of utmost importance. For O'Shaughnessy is claiming that bodily-awareness is a transcendental condition of action, yet these remarks suggest that this might not be so. Clearly, if it is true that immediate presence is required for immediately moving a limb, then a creature entirely lacking bodily-awareness would not be able to partake in the first kind of instrumental action. It would not be able to move its right arm with its left arm since it would not be able to move its left arm. But might not such a creature be able to instrumentally move its limbs via 'a mental act like visualising'? Maybe the following is a possible situation: the creature lacks bodily-awareness but is capable of engaging in the basic (non-spatial) act of 'visualising'. By doing this the creature can effect instrumental movements of its limbs. Thus, this creature whilst lacking bodily-awareness is nevertheless capable of intentional action.

Not only does this seem perfectly conceivable, there actually exist people who are like this. As mentioned in chapter seven, there have been several well documented cases of people who have gone 'body-blind'. These are, effectively, people who lack bodily-awareness as we know it. They are, however, still capable of intentional

---

<sup>20</sup> See Danto, 1965.

bodily-actions. Although, in all cases, bodily-action was not possible immediately after the onsets of their various neuropathies, they gradually learned to perform quite advanced movements including, for example, walking.

This appears to be a straightforward counterexample to the claim that bodily-awareness is a transcendental condition of spatial action. Whilst bodily-awareness is, in normal subjects, very intimately involved in the ability to perform spatial actions, it is not strictly necessary. It is clear that when a normal subject raises their right arm, say, they do so partly on the basis of their awareness of where their arm is. This is so natural as to be frequently overlooked. But, as the cases of body-blindness show, spatial actions can be performed in the absence of this awareness. But how do body-blind patients manage to act? Apparently, it seems that vision (and maybe hearing) and intense concentration replace bodily-awareness as the basis for action,

Constant visual vigilance is required for any purposeful movement....this visual feedback can only be used with concentration and intellectual effort. (Cole & Paillard, 1995, p.250)

This suggests that body-blind patients manage to perform spatial actions *via* something akin to O'Shaughnessy's 'mental act like visualising'. If this is correct, then such movements could hardly be classed as basic actions, but they are actions nevertheless. It seems, then, that bodily-awareness is not a transcendental condition of spatial action. And this means that the experience-directed argument fails<sup>21</sup>. We have been given no reason to think it impossible for a subject, who has never had experience of a body as his, to engage in spatial action. For we have been given, in the case of body-blindness, a model upon which to base our understanding of how such a subject might go about this. Such a subject would be in a situation radically unlike our own, but we should not doubt its coherence<sup>22</sup>.

---

<sup>21</sup> Another reason for doubting O'Shaughnessy's claim concerning bodily-awareness and spatial action comes in the form of certain experimental data. In an experiment described in Marcel, 2003, pp.62-7, subject's right arms were occluded under a surface which had a light on it. Due to the muscle tendon of the arm being vibrated, subjects experienced their arm to be left of the light when it was actually to the right. That is, they suffered a 'vibro-tactile illusion'. The subjects were then asked to move their hand to underneath the light. Despite the illusion, they had no difficulty doing this, their arm moving leftwards. However, subjects reported that they felt that they had moved their arms to the right. Here, whilst the subjects awareness of their action is illusory, their execution of the action is unimpaired. This is relevant to the current discussion in that it would appear to drive a wedge between feeling a limb to be present and positioned in a certain way, and being able to act intentionally with it.

<sup>22</sup> Such a subject would, perhaps, be like Descartes reminds me that I am not, "present in my body as a sailor is present in a ship" (Descartes, 1996, p.56).

## 8.7 Having a Body

Recall that the action argument is attempting to show that Cassam's second grade of apparent presence in the world is a transcendental condition of self-consciousness. This grade supposedly involves an awareness of oneself as 'having' a body. There are two ways in which this might be interpreted. The first is that one must experience oneself as having a body. The considerations in the previous section tell against this. But there is another possible interpretation. Think again of body-blind people. People suffering from body-blindness can *see* their bodies (they manage to act via visual monitoring). They still make bodily self-ascriptions. Such people can look at their arm and make the visually-based judgement that their arm is bent, or whatever. Body blind people still have bodies which they continue to regard as theirs. Although they do not *feel* their bodies as theirs, they still *think* of their bodies as theirs. Maybe this is the key. We can now state the main contention of the belief-directed action argument. It is this: if a subject neither felt, nor thought of, a body as theirs they would not be able to engage in intentional spatial action. Here, the second grade of apparent presence in the world is being read as *regarding* oneself as having a body.

We all 'have' bodies, no non-sceptic will deny this. Even the most ardent dualist will allow that we bear a relation to our bodies that we bear to nothing else. My body is mine, not like my record player is mine, but in some far more primitive sense. This sense of mineness, or ownership, is most obviously apparent in bodily-awareness. For, as was claimed in chapter four, in bodily-awareness we *feel* this sense of ownership. As Martin has put it,

When I feel an ache in my ankle, the ankle that feels hurt to me does not just feel like an ankle belonging to some body or other. Rather the ankle feels to me to be part of my body (Martin, 1995, p.269)

Cases of body-blindness show that such a felt sense of ownership is not a necessary condition of spatial action, for body-blind subjects lack such a sense. But there is a cognitive parallel. As well as feeling our bodies to be ours, we consider, or regard, our bodies to be ours. And body-blind subjects continue to do so. We can understand this 'regarding a body to be one's own' as being closely related to the disposition to self-ascribe bodily predicates.

If a subject no longer felt or regarded his body as his, could he perform spatial actions? It is not obvious how. For such a subject would have no reason to try to move his arm rather than, say, a nearby table. If an already body-blind person were to suffer another neuropathy, the consequence of which was that he no longer regarded his body as any more part of him than any other physical object he would, on this suggestion, be incapable to spatial action.

To this claim it will be objected that if a body-blind subject can instrumentally move his arm by some kind of mental act, then the currently envisaged subject could do much the same thing. So, the body-blind subject performs a mental act of visualising his arm moving, and by doing this he can move his arm. Mightn't it be that, in a similar way, a subject who does not regard his body as his could perform the mental act of visualising *that* arm as moving and thereby move it? After all, we can perform instrumental actions with objects which we do not consider 'ours' – I open the can with the can opener. If such a case is coherent, haven't we described a subject who is capable of spatial action (albeit only instrumentally) but who does not satisfy the conditions of the second grade of apparent presence in the world? Two responses suggest themselves to this objection. The first is to claim that, despite first appearances, this does not represent a real possibility. The second is to claim that although the envisaged situation is possible, it is not true to say that such a subject does not think of a body as his. I shall discuss these in turn.

One reason for thinking that a subject who did not feel, or even regard, their body as theirs would be incapable of (instrumental) spatial action is that such a subject would not understand the request 'Move *that* arm'. Consider for a moment an ordinary subject. If we say to him 'Move that table' he will walk over to it and push. If we say 'No, move the table in the same way as you move your arm', he will be at a loss<sup>23</sup>. For if we are asked to will a movement in an extra-bodily object we don't even know where to begin. Do we think 'Move!', or mutter it under our breath? It is plausible to think that it just doesn't make sense to ask this of someone. But isn't our subject in much the same position with regards to his limbs? Since he does not regard them as his, he is in the same position regarding them as we are to tables. Thus the envisaged situation is not, in fact, possible.

---

<sup>23</sup> This is in line with the general thrust of O'Shaughnessy, 1980, Chs.2-4.

But this line of thought is flawed. For what we (arguably) don't understand is how to move a table immediately; as a basic action. We can perfectly well move it as an instrumental action. We have already allowed that the subject in our example cannot move his limbs immediately; as basic actions. The question is whether he can move them instrumentally. Since it is perfectly easy for us to understand how it is possible for us to move extra-bodily objects instrumentally, we have no reason to think that the subject who did not feel, or even regard, their body as theirs would be incapable of moving their limbs instrumentally.

What about the other response? The claim that although the envisaged situation is possible, it is not true to say that such a subject fails to think of a body as his? Let's review the situation. We have a subject who has no feeling in, or sense of his own body; he is body-blind. Furthermore, he does not regard his body to be his. That is, he does not self-ascribe any bodily predicates. He can see, what is in fact, his arm before him but does not think of this arm as 'his', any more than he thinks of this chair before him as 'his'. Now we add that this subject can perform mental actions of visualisation, via which he can instrumentally move what is in fact his arm. Imagine that the subject does this, and begins to instrumentally act with his body. Now after a while the subject is bound to notice that it is much easier to move one object than it is to move any others. For one object (his body) can be moved simply by visualising it. Whereas other objects can only be moved by visualising that one object as moving them. Thus, he will notice that he bears some sort of special relationship to that object. So, any subject in this situation has the capacity to come to realise that he bears a special relationship to one physical object, viz. it is the only object which he can move simply by visualising it. Isn't that subject now thinking of that object which is in fact his body as 'his'? Since this subject is ex hypothesi a self-ascriber, isn't he now in a position to self-ascribe bodily predicates?

He may well be, but that doesn't mean that he will. The test that we should use for whether a subject regards a body as his is whether or not he is disposed to make bodily self-ascriptions. But the fact, if it is one, that the subject currently under discussion is in a position to make such self-ascriptions does not show that he would be disposed to do so. It is still perfectly conceivable that this subject regard the object with which he has this special relationship as just that. That is, as just some object with which he has a special relationship as regards spatial action.

Once we are in this position, it becomes difficult to find a convincing reason to reject an even more extreme possibility. The possibility of Barny the angel. If the following is a possibility, then the belief-directed version of the action argument must be rejected. Barny is an angel. As far as he can tell he has no body, but is a purely spiritual entity<sup>24</sup>. Barny is a ‘point of view’ on the world<sup>25</sup>. We might think of Barny on the model of an invisible man, suffering from congenital body-blindness. But Barny can perform spatial actions. These actions take the form of moving backwards and forwards, left and right, up and down, and of rotating clockwise or anti-clockwise. He simply wills the movement of his point of view, and it occurs. Or maybe, he performs a mental action of visualising the movement, and then it occurs as a result.

In this example Barny does not have some hidden disposition to self-ascribe bodily predicates. It looks as though the defender of the action argument is going to have to bite the bullet here and argue that Barny is not a possibility. How might this be done? Maybe it could be argued that the example is altogether too magical. It might be argued that for a spatial action to occur, there must be some sort of physical mechanism in place which partially explains the occurrence of the movement<sup>26</sup>. In the case of Barny, there is no such mechanism and therefore no action. But why assume that there is no mechanism? Perhaps Barny sits on an invisible, or glass platform. This cannot be perceived by Barny, but it is surely there. Further, there is an intricate mechanism linking Barny’s (invisible) brain with his movable glass platform, such that willing it to move forward makes it move forward. Now we have a mechanism, but Barny still will not be disposed to self-ascribe any bodily predicates. The absence of mechanism doesn’t seem to get us where we want to go. It doesn’t show the impossibility of Barny’s performing spatial actions.

Some philosophers think that the situation I have described concerning Barny and his willed movements is, in fact, incoherent. For instance, McDowell suggests that when we consider Barny, we are considering “a self that recedes inward to the point of vanishing. With bodily agency in the picture, that bodily thing becomes something one can identify as oneself; now there is plainly something present in the

---

<sup>24</sup> This is consistent with him actually having a body. The point is that if he does have a body, he can neither feel, see, smell, hear or taste it.

<sup>25</sup> This possibility is akin to what Cassam calls ‘geometrical self-location’, Cassam, 1997, pp.44-51. Also, see McDowell, 1994, pp.99-104.

<sup>26</sup> See, O’Shaughnessy, 1980, Ch. 4.

world for oneself to be” (McDowell, 1998, p.143)<sup>27</sup>. It is true that there is something difficult to grasp about the possibility of the will being exercised on something as insubstantial as a point of view. But the fact that it is difficult to grasp does not show that it is strictly inconceivable. Indeed, the point of the present section has been to suggest that we can, in fact, conceive of such a situation. That considerations of the nature of spatial perception, and spatial action do not serve to show that Barny is impossible (that a self-conscious subject must satisfy the second grade of apparent presence in the world). So, on the basis of the action argument, we cannot conclude that embodied experience is a transcendental condition of self-consciousness.

---

<sup>27</sup> Also see, McDowell, 1994, pp.99-104.

## 9. Conclusions

My official position towards the Body Claim is agnostic, and I hope to have shown that influential arguments in its favour are wanting. Whilst this conclusion is a negative one, we should not lose sight of the many positive claims that have been defended in the process. Although the Body Claim has not been substantiated, much has been learned along the way about both self-consciousness and embodied experience.

In the first three chapters I defended the claim that central cases of self-consciousness necessarily involve self-reference. It was pointed out that arguments in favour of the Body Claim in fact tend to offer transcendental conditions on self-reference. For this to be relevant to self-consciousness, it must to be shown that self-consciousness necessarily involved self-reference. I argued against the two most serious objections to this view. Both the no reference view and Evans' theory of singular thought pose *prima facie* challenges to the connection between self-consciousness and self-reference. Those challenges, however, are not successful. The chapters that followed have been an examination of some of the consequences that might be thought to follow from the idea that self-consciousness necessarily involves self-reference.

First, I argued against the view that an introspective awareness of oneself as a bodily subject is a transcendental condition on self-consciousness. We ourselves are counterexamples to this claim, since we are self-conscious, and yet do not have an introspective awareness of ourselves as bodily-subjects. Second, I argued that experiencing oneself or thinking of oneself as having a body was not, strictly speaking, a transcendental condition of self-consciousness either. This denial is different, since we evidently do (in the usual case) both experience ourselves and think of ourselves as having bodies. I have defended the view that self-consciousness requires that one think of oneself, and experience oneself, as spatially located. From the solidity argument I have concluded that (given certain empirical assumptions) spatial experience requires one to be able to think about internally causally connected things. From the action argument, I concluded that a self-conscious subject must also have some grasp of the possibility of its spatially acting. These are, of themselves, philosophically interesting conclusions, even though they do not achieve what the solidity and action arguments intended.

Central to the thesis has been the denial of Russell's Principle. One of the most important morals is that the defender of the Body Claim via the notion of self-reference needs to pay especial notice of his or her relation to Russell's principle. If Russell's Principle is accepted, it is difficult to forge the necessary connection between self-consciousness and self-reference. If Russell's Principle is rejected, many of the most popular arguments for the Body Claim are undermined.

## 10. Appendix

Pryor, 1999, defines immunity to error in the following way (he refers to it as ‘immunity to *de re* misidentification’),

a proposition is **immune to *de re* misidentification when justified by grounds G** just in case, whenever that proposition is believed on grounds G, then the resulting belief can not be in error through *de re* misidentification (Pryor, 1999, p.279)

Of course, we now need a definition of *de re* misidentification. Pryor says that we have a case of *de re* misidentification whenever the following three conditions are satisfied:

- i. There is some singular proposition about  $x$ , to the effect that it is  $F$ , that a subject believes or attempts to express.
- ii. The subject’s justification for believing this singular proposition rests on his justification for believing, of some  $y$ , and of  $x$ , that  $y$  is  $F$  and that  $y$  is identical to  $x$ .
- iii. Unbeknownst to the subject,  $y \neq x$ .

Given this account of errors of misidentification, we can formulate a definition of immunity to error as follows:

**A belief of the form ‘ $a$  is  $F$ ’ is *de re* immune when based on grounds G iff the following is not possible:**

- (1) S’s justification for believing that  $a$  is  $F$  rests on his justification for believing of some  $b$  that it is  $F$  and his justification for believing that  $a$  is identical to  $b$ .**
- (2)  $a$  is not identical to  $b$ .**

Notice that this definition does not require  $S$  to believe either that  $b$  is  $F$  or that  $a$  is identical to  $b$ . So, it cannot require  $S$ ’s belief that  $a$  is  $F$  to be *based* on these two beliefs. Rather, the justification for  $S$ ’s belief that  $a$  is  $F$  rests on his justification for

believing these two propositions. However, it does seem to follow that S is *justified* in believing that *b* is *F* and that *a* is identical to *b*, for assuming that S is justified in believing that *a* is *F*, this justification could not rest on his justification for believing that *b* is *F* and that *a* is identical to *b*, unless he were justified in believing the second two propositions. On the other hand, if S is not justified in believing that *a* is *F*, his non-existent justification cannot rest on anything.

Therefore, in one respect, Pryor's definition of *de re* misidentification (not *de re* immunity) is weaker than that implicit in the standard version (see chapter two), in that S need not actually believe that *b* is *F* or that *a* is identical to *b*.

I have two objections to Pryor's way of formulating IEM, one technical and one rather more general. The first, technical, problem is this<sup>1</sup>. The right-hand side of the biconditional can be formalized as, ' $\neg\Diamond(1 \wedge 2)$ ', which is equivalent to, ' $\neg\neg\Box\neg(1 \wedge 2)$ '. But that is equivalent to, ' $\Box\neg(1 \wedge 2)$ ', which is equivalent to, ' $\Box(\neg 1 \vee \neg 2)$ '. However, since ' $a=b \vdash \Box(a=b)$ ' is true, it follows that ' $\neg 2 \vdash \Box\neg 2$ ' is true.

So, the RHS will be true whenever  $a=b$ . But, this is the wrong result. It entails that a proposition is IEM whenever it can be based on a true identity judgement. So that 'Hesperus is a star' will be IEM since it can be based on the judgements 'Phosphorus is a star', and 'Hesperus=Phosphorus'.

This difficulty can be circumvented, however, by altering the definition. I suggest that the notion Pryor is after can be captured in the following,

**The proposition '*a* is *F*' is immune to error relative to grounds *G* iff the justification for every possible judgement that *a* is *F* (based on grounds *G*) either:**

**(1\*) does not rest on the justification for two judgements of the form '*b* is *F*' and '*a=b*'.**

**or,**

**(2\*) rests on the justification for two judgements of the form '*b* is *F*' and '*a=b*', but '*a=b*' is true *a priori*.**

---

<sup>1</sup> I thank Nick Shackel for pointing out and patiently explaining to me, a similar problem with an earlier definition of IEM I myself offered.

Like Pryor's definition, this definition does not require S to believe either that *b* is *F* or that *a* is identical to *b*. So, again, there is a sense in which the notion of misidentification employed in this definition is weaker than that employed in the definition of IEM given in chapter two. This point is important when it comes to my second, more general, objection to Pryor.

Since there is a respect in which Pryor's definition (of misidentification) is weaker than the standard definition, there should be some cases which qualify as instances of *de re* misidentification under Pryor's definition but not under the standard account. That is, we should be able to give an example of a subject who believes that *a* is *F* and who, whilst he is *justified* in believing both that some *b* is *F* and that *a* is identical to *b*, he doesn't actually hold these two beliefs. Further, his justification for believing that *a* is *F* rests on his justification for believing these two propositions. However, unbeknownst to the subject, *a* is not identical to *b*. Unfortunately, none of the six examples of *de re* misidentification Pryor gives in his paper help us here, since in all of them the subject does actually believe that *b* is *F* and that *a* is identical to *b*. Neither have I managed to come with an example. In fact, I'm not entirely sure an example is there to be thought of. To see why consider the following case.

Someone who looks just like Bert is standing in front of me. He's wearing blue socks. Since I can see this person I am justified in believing that the person in front of me is wearing blue socks and that the person in front of me is Bert. I am also justified in believing that Bert is wearing blue socks and this justification rests on my justification for believing that the person in front of me is wearing blue socks and that the person in front of me is Bert. However, I have the groundless belief that I am hallucinating and so I never actually form the belief that the person in front of me is wearing blue socks or the belief that the person in front of me is Bert. But, of course, it is not Bert in front of me, it's Ernie. Importantly, this is not yet a case of *de re* misidentification. In order for it to be so, it must be the case that I do actually believe that Bert is wearing blue socks. But it seems to me wholly implausible to suggest that this can simply be added to the example. Why do I believe that Bert is wearing blue socks? Surely not because of my visual perceptions, since I consider them to be illusory. Some other reason then, say a reliable source tells me that Bert is wearing blue socks. But then it seems unlikely that (ii) will continue to hold. If I believe that Bert is wearing blue socks on this new evidence, it seems likely that my *justification* for believing it will rest upon my justification for believing the testimony not upon my

justification for believing that the person standing in front of me is wearing blue socks and that the person standing in front of me is Bert. But if this is right, Pryor's definition will collapse into a version of the standard account. Any belief which qualifies as an instance of *de re* misidentification in Pryor's sense, will qualify as an instance of misidentification in the standard sense. As a result, I favour the definition of IEM given in chapter two.

## References

[Where applicable, page references are to reprinted versions]

- Alexander, P. 1991. Locke on Solidity and Incompressibility. In Mahalingam and Carr, 1991, pp.149-157.
- Almog, J., Perry, J. and Wettstein, H. (eds.). 1989. *Themes From Kaplan*. Oxford: Oxford University Press.
- Anscombe, G. E. M. 1975. The First Person. In Guttenplan, 1975, pp.45-65.
- Aquila, R. E. 1979. Personal Identity and Kant's "Refutation of Idealism". *Kant-Studien* 70, pp.259-78.
- Austin, J. L. 1962. *How to Do Things With Words*. Oxford: Clarendon Press.
- Ayers, M. 1991. *Locke*, 2 vols. London: Routledge.
- Bach, K. 1987. *Thought and Reference*. Oxford: Clarendon Press.
- Bar-On, D. 2000. Speaking My Mind. *Philosophical Topics* 28, pp.1-34.
- Bar-On, D. and Long, D. C. 2001. Avowals and First-Person Privilege. *Philosophy and Phenomenological Research* 62, pp.311-335.
- Beck, L. W. (ed.). 1969. *Kant Studies Today*. La Salle, Ill.: Open Court.
- Bennett, J. 1966. *Kant's Analytic*. Cambridge: Cambridge University Press.
- Bennett, J. 1968. Strawson on Kant. *The Philosophical Review* 77, pp.340-349.
- Bennett, J. 1971. *Locke, Berkeley, Hume: Central Themes*. Oxford: Clarendon Press.
- Bermúdez, J. L. 1995. Transcendental Arguments and Psychology: The Example of O'Shaughnessy on Intentional Action. *Metaphilosophy* 26, pp.379-401.
- Bermúdez, J. L. 1998. *The Paradox of Self-Consciousness*. Cambridge, Mass.: MIT Press.
- Bermúdez, J. L., Marcel, A., and Eilan, N. (eds.). 1995. *The Body and the Self*. Cambridge, Mass.: MIT Press.
- Blackburn, S. 1984. *Spreading the Word: Groundings in the Philosophy of Language*. Oxford: Clarendon Press.
- Bloom, P., Peterson, M. A., Nadel, L. and Garrett, M. F. (eds.). 1996. *Language and Space*. Cambridge, Massachusetts: MIT.

- Brandom, R. B. 1994. *Making It Explicit: Reasoning, Representing, and Discursive Commitment*. Cambridge, Mass.: Harvard University Press.
- Brentano, F. 1973. *Psychology From an Empirical Standpoint*. Translated by Rancurello, Terrell and McAlister. London: Routledge.
- Brewer, B. 1992. Self-Location and Agency. *Mind* 101, pp.17-34.
- Brewer, B. 1995. Bodily Awareness and the Self. In Bermúdez, Marcel and Eilan, 1995, pp.291-309.
- Brinck, I. 1997. *The Indexical 'I'*. Dordrecht: Kluwer Academic Publishers.
- Bryant, D. J. 1997. Representing Space in Language and Perception. *Mind & Language* 12, pp.239-264.
- Burge, T. 1998. Reason and the First Person. Wright, Smith and MacDonald, 1998, pp.243-270.
- Burke, M. B. 1994. Preserving the Principle of One Object to a Place: A Novel Account of the Relations Among Objects, Sorts, Sortals and Persistence Conditions. *Philosophy and Phenomenological Research* 54, pp.591-624.
- Campbell, J. 1993. The Role of Physical Objects in Spatial Thinking. In Eilan, McCarthy and Brewer, 1993, pp.65-95.
- Campbell, J. 1994a. *Past, Space, and Self*. Cambridge, Mass.: MIT Press.
- Campbell, J. 1994b. Objects and Objectivity. In Peacocke, 1994, pp.3-20.
- Campbell, J. 1997. Replies. *Philosophy and Phenomenological Research* 62, pp.655-670.
- Campbell, J. 1999a. Immunity to Error through Misidentification and the Meaning of a Referring Term. *Philosophical Topics* 26, pp.89-104.
- Campbell, J. 1999b. Schizophrenia, the Space of Reasons, and Thinking as a Motor Process. *The Monist* 82, pp.609-625.
- Campbell, J. 2002. *Reference and Consciousness*. Oxford: Clarendon Press.
- Cassam, Q. 1994. *Self-Knowledge*. Oxford: Oxford University Press.
- Cassam, Q. 1995a. Introspection and Bodily Self-Ascription. In Bermúdez, Marcel and Eilan, 1995, pp.311-36.
- Cassam, Q. 1997. *Self and World*. Oxford: Oxford University Press.
- Cassam, Q. 2002. Representing Bodies. *Ratio (New Series)* 15, 315-334.

- Castañeda, C. 1966. He: A Study in the Logic of Self-Consciousness. *Ratio* 8. Reprinted in Castañeda, 1999, pp.35-60.
- Castañeda, C. 1999. *The Phenomeno-Logic of the I: Essays on Self-Consciousness*. Edited by J. G. Hart & T. Kapitan. Bloomington and Indianapolis: Indiana University Press.
- Chisholm, R. 1969. On the Observability of the Self. *Philosophy and Phenomenological Research* 30. Reprinted in Cassam, 1994, pp.94-108.
- Clarke, D. S. 1978. The Addressing Function of 'I'. *Analysis* 38, pp.91-3.
- Cole, J. 1991. *Pride and a Daily Marathon*. London: Duckworth.
- Cole, J. and Paillard, J. 1995. Living Without Touch and Peripheral Information about Body Position and Movement: Studies with Deafferented Subjects. In Bermúdez, Marcel and Eilan, 1995, pp.245-266.
- Corazza, E., Fish, W., and Gorvett, J. 2002. Who is I?. *Philosophical Studies* 107, pp.1-21.
- Corcoran, K. (ed.). 2001. *Soul, Body, and Survival: Essays on the Metaphysics of Human Persons*. Ithaca: Cornell University Press.
- Craig, E. 1987. *The Mind of God and the Works of Man*. Oxford: Clarendon Press.
- Crane, T. (ed.). 1992. *The Contents of Experience*. Cambridge: Cambridge University Press.
- Crowley, C. J. 1980. *Strawson's Theory of the Person in Individuals*. Ann Arbor, Michigan: University Microfilms International.
- Dainton, B. 2000. *Stream of Consciousness*. London: Routledge.
- Danto, A. 1965. Basic Actions. *American Philosophical Quarterly* 2. Reprinted in White, 1968, pp.43-58.
- Davidson, D. 1968. On Saying That. *Synthese* 19. Reprinted in Davidson, 1984, pp.93-108.
- Davidson, D. 1984. *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press.
- de Gaynesford, M. 2002. Corporeal Objects and the Interdependence of Perception and Action. *Ratio (New Series)* 15, pp.335-353.
- Descartes, R. 1996. *Meditations on First Philosophy*. Translated and edited by J. Cottingham. Cambridge: Cambridge University Press.
- Diamond, C. and Teichman, J. (eds.). 1979. *Intention and Intentionality: Essays in Honour of G. E. M. Anscombe*. Brighton: Harvester Press.

- Eilan, N. 1995. Consciousness and the Self. In Bermúdez, Marcel and Eilan, 1995, pp.337-357.
- Eilan, N. 1997. Objectivity and the Perspective of Consciousness. *European Journal of Philosophy* 5, pp.235-250.
- Eilan, N., McCarthy, R. and Brewer, B. (eds.). 1993. *Spatial Representation: Problems in Philosophy and Psychology*. Oxford: Oxford University Press.
- Evans, G. 1980. Things Without the Mind – A Commentary upon Chapter Two of Strawson's *Individuals*. In Van Straaten, 1980, pp.76-116.
- Evans, G. 1982. *The Varieties of Reference*. Edited by J. McDowell. Oxford: Clarendon Press.
- Falvey, K. 2000. The Basis of First-Person Authority. *Philosophical Topics* 28, pp.69-99.
- Freud, S. 1927. The Ego and the Id. In Freud, 1991, pp.339-407.
- Freud, S. 1991. *On Metapsychology*. Compiled and edited by A. Richards. London: Penguin.
- Fricke, M. F. and Snowdon, P. Forthcoming. Solidity and Impediment. *Analysis* 63.
- Gallagher, S. 2000. Self-Reference and Schizophrenia. In Zahavi, 2003, pp.203-231.
- Gallagher, S. and Shear, J. (eds.). 1999. *Models of the Self*. Exeter: Imprint Academic.
- Gardner, S. 1994. Other Minds and Embodiment. *Proceedings of the Aristotelian Society* 94, pp.35-52.
- Geach, P. T. 1957. *Mental Acts: Their content and their objects*. London: Routledge.
- Geach, P. T. 1965. Assertion. *Philosophical Review* 74. Reprinted in Geach, 1972, pp.254-269.
- Geach, P. T. 1972. *Logic Matters*. Oxford: Blackwell.
- Geach, P. T. 1986. Review of Evans, *The Varieties of Reference*. *Philosophy* 61, pp.534-8.
- Gendler, Z. S. and Hawthorne, J. (eds.). 2002. *Conceivability and Possibility*. Oxford: Clarendon Press.
- Glock, H-J. and Hacker, P. M. S. 1996. Reference and the First Person Pronoun. *Language and Communication* 16, pp.95-105.
- Guttenplan, S. (ed.). 1975. *Mind and Language*. Oxford: Clarendon Press.

- Hacker, P. M. S. 1990. *Wittgenstein: Meaning and Mind, An Analytical Commentary on the Philosophical Investigations, vol. III*. Oxford: Blackwell.
- Hahn, L. E. (ed.). 1998. *The Philosophy of P. F. Strawson*. Chicago: Open Court.
- Haldane, J. and Wright, C. (eds.). 1993. *Reality, Representation and Projection*. New York: Oxford University Press.
- Hale, B. 1993. Can there be a Logic of Attitudes?. In Haldane and Wright, 1993, pp.337-363.
- Hamilton, A. 1991. Anscombian and Cartesian Scepticism. *The Philosophical Quarterly* 41, pp.39-54.
- Hamilton, A. 1995. A New Look at Personal Identity. *The Philosophical Quarterly* 45, pp.332-349.
- Hampshire, S. 1982. *Thought and Action*. New Edition. London: Chatto and Windus.
- Harcourt, E. 2000. The First-Person: Problems of Sense and Reference. In Teichmann, 2000, pp.25-46.
- Harrison, R. 1970. Strawson on Outer Objects. *The Philosophical Quarterly* 20, pp.213-221.
- Heal, J. 1994. Moore's Paradox: A Wittgensteinian Approach. *Mind* 103, pp.5-24.
- Heidegger, M. 1962. *Being and Time*. Translated by J. Macquarrie & E. Robinson. Oxford: Blackwell.
- Henrich, D. 1989. The Identity of the Subject in the Transcendental Deduction. In Schaper and Vossenkuhl, 1989, pp.250-80.
- Hume, D. 1978. *A Treatise of Human Nature*. Edited by L. A. Selby-Bigge, revised by P. H. Nidditch. Oxford: Oxford University Press.
- Hurley, S. 1994. Unity and Objectivity. In Peacocke, 1994, pp.49-77.
- Ishiguro, H. 1980. The Primitiveness of the Concept of a Person. In Van Straaten, 1980, pp.62-75.
- Jacobsen, R. 1996. Wittgenstein on Self-Knowledge and Self-Expression. *The Philosophical Quarterly* 46, pp.12-30.
- Janaway, C. 1989. *Self and World in Schopenhauer's Philosophy*. Oxford: Clarendon Press.
- Joske, W. D. 1967. *Material Objects*. London: Macmillan.
- Kalderon, M. E. Forthcoming. *Moral Fictionalism*.

- Kant, I. 1997. *The Critique of Pure Reason*. Translated by Guyer & Wood. Cambridge: Cambridge University Press.
- Kaplan, D. 1977. Demonstratives. In Almog, Perry and Wettstein, 1989, pp.481-563.
- Katz, D. 1989. *The World of Touch*, edited and translated by L. E. Krueger. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.
- Kenny, A. 1979. The First Person. In Diamond and Teichman, 1979, pp.3-13.
- Körner, S. 1969. The impossibility of Transcendental Deductions. In Beck, 1969, pp.230-244.
- Kripke, S. 1981. *Naming and Necessity*. Oxford: Blackwell.
- Kvart, I. 1989. Divided Reference. *Midwest Studies in Philosophy* 14, pp.140-179.
- Levinson, S. C. 1996. Frames of Reference and Molyneux's Question: Crosslinguistic Evidence. In Bloom, Peterson, Nadel and Garrett, 1996, pp.109-169.
- Lewis, D. 1970. General Semantics. *Synthese* 22. Reprinted in Lewis, 1983, pp.189-232.
- Lewis, D. 1976. Survival and Identity. In Rorty, 1976. Reprinted with 'Postscripts to "Survival and Identity"' in Lewis, 1983, pp.55-77.
- Lewis, D. 1979. Attitudes *De Dicto* and *De Se*. *The Philosophical Review* 88. Reprinted with 'Postscripts to "Attitudes *De Dicto* and *De Se*"' in Lewis, 1983, pp.133-159.
- Lewis, D. 1983. *Philosophical Papers, Vol. I*. New York: Oxford University Press.
- Loar, B. 1984. Review of *The Varieties of Reference*. *Philosophical Books* 25, pp.46-51.
- Locke, J. 1975. *An Essay Concerning Human Understanding*, edited by P. H. Nidditch. Oxford: Clarendon.
- Mahalingam, I. and Carr, B. (eds.). 1991. *Logical Foundations: Essays in Honour of D. J. O'Connor*. Basingstoke: Macmillan.
- Malcolm, N. 1979. Whether 'I' is a Referring Expression. In Diamond and Teichman, 1979, pp.15-24.
- Marcel, A. 2003. The Sense of Agency: Awareness and Ownership of Action. In Roessler and Eilan, 2003, pp. 48-93.
- Martin, M. G. F. 1992. Sight and Touch. In Crane, 1992, pp.196-215.

- Martin, M. G. F. 1993. Sense Modalities and Spatial Properties. In Eilan, McCarthy and Brewer, 1993, pp.206-218
- Martin, M. G. F. 1995. Bodily-Awareness: A Sense of Ownership. In Bermúdez, Marcel and Eilan, 1995, pp.267-289.
- Martin, M. G. F. 1997. Self-Observation. *European Journal of Philosophy* 5, 119-40.
- McDowell, J. 1990. Peacocke and Evans on Demonstrative Content. *Mind* 99, pp.255-266.
- McDowell, J. 1994. *Mind and World*. Cambridge, Mass.: Harvard University Press.
- McDowell, J. 1998. Referring to Oneself. In Hahn, 1998, pp.129-145.
- McGinn, C. 1982. *The Character of Mind: An Introduction to the Philosophy of Mind*. Oxford: Oxford University Press.
- Merleau-Ponty, M. 1962. *Phenomenology of Perception*. Translated by C. Smith. London: Routledge.
- Millikan, R. G. 2000. *On Clear and Confused Ideas*. Cambridge: Cambridge university Press.
- Nagel, T. 1971. Brain Bisection and the Unity of Consciousness. *Synthese* 20. Reprinted in Nagel, 1979, pp.147-164.
- Nagel, T. 1979. *Mortal Questions*. Cambridge: Cambridge University Press.
- O'Brien. L. 1994. Anscombe and the Self-Reference Rule. *Analysis* 54, pp.277-281.
- O'Brien, L. 1995. Evans on Self-Identification. *Noûs* 29, pp.232-247.
- Olson, E. T. 2001. A Compound of Two Substances. In Corcoran, 2001, pp.73-88.
- O'Shaughnessy, B. 1980. *The Will: A Dual Aspect Theory*, 2 vols. Cambridge: Cambridge University Press.
- O'Shaughnessy, B. 1989. The Sense of Touch. *Australasian Journal of Philosophy* 67, pp.37-53.
- O'Shaughnessy, B. 1995. Proprioception and the Body Image. In Bermúdez, Marcel and Eilan, 1995, pp.175-203.
- O'Shaughnessy, B. 2000. *Consciousness and the World*. Oxford: Clarendon Press.
- Paillard, J. (ed.). 1991. *Brain and Space*. Oxford: Oxford University Press.

- Peacocke, A. and Gillett, G. (eds.). 1987. *Persons and Personality: A Contemporary Inquiry*. Oxford: Blackwell.
- Peacocke, C. 1983. *Sense and Content: Experience, Thought and their Relations*. Oxford: Clarendon Press.
- Peacocke, C. 1989. *Transcendental Arguments in the Theory of Content: An Inaugural Lecture delivered before the University of Oxford on 16 May 1989*. Oxford: Clarendon Press.
- Peacocke, C. 1991. Demonstrative Content: A Reply to John McDowell. *Mind* 100, pp.123-133.
- Peacocke, C. 1992. Scenarios, Concepts and Perception. In Crane, 1992, pp.105-135.
- Peacocke, C. 1993. Intuitive Mechanics, Psychological Reality and the Idea of a Material Object. In Eilan, McCarthy and Brewer, 1993, pp.162-176.
- Peacocke, C. 1999. *Being Known*. Oxford: Clarendon Press.
- Peacocke, C. 2003. Action: Awareness, Ownership and Knowledge. In Roessler and Eilan, 2003, pp.94-110.
- Peacocke, C. (ed.). 1994. *Objectivity, Simulation and the Unity of Consciousness: Current Issues in the Philosophy of Mind*. Oxford: Oxford University Press.
- Penelhum, T. 2000. *Themes in Hume*. Oxford: Clarendon Press.
- Perry, J. 1977. Frege on Demonstratives. *The Philosophical Review* 86. Reprinted in Perry, 2000, pp.1-26.
- Perry, J. 1979. The Problem of the Essential Indexical. *Noûs* 13. Reprinted in Perry, 2000, pp.27-44.
- Perry, J. 2000. *The Problem of the Essential Indexical and Other Essays*. Expanded Edition. Stanford: CLSI Publications.
- Pitson, A. E. 2002. *Hume's Philosophy of the Self*. London: Routledge.
- Pryor, J. 1999. Immunity to Error through Misidentification. *Philosophical Topics* 26, pp.271-304.
- Quinton, A. 1962. Spaces and Times. *Philosophy* 37, pp.130-147.
- Quinton, A. 1964. Matter and Space. *Mind* 73, pp.332-352.
- Rea, M. (ed.). 1997. *Material Constitution: A Reader*. Lanham, Maryland: Rowman & Littlefield.

- Roessler, J. and Eilan, N. (eds.). 2003. *Agency and Self-Awareness: Issues in Philosophy and Psychology*. Oxford: Clarendon Press.
- Rorty, A. O. (ed.). 1976. *The Identities of Persons*. Berkley: University of California Press.
- Rorty, R. 1970. Strawson's Objectivity Argument. *Review of Metaphysics* 24, pp.207-244.
- Rozemond, M. 1992-3. Evans on *De Re* Thought. *Philosophia* 22, pp.275-298.
- Russell, B. 1910-11. Knowledge by Acquaintance and Knowledge by Description. *Proceedings of the Aristotelian Society* 11. Reprinted in Russell, 1917, pp.152-167.
- Russell, B. 1917. *Mysticism and Logic*. London: Unwin.
- Russell, B. 1921. *The Analysis of Mind*. London: Unwin.
- Russell, B. 1956. *Logic and Knowledge*. London: Unwin.
- Sacks, M. 2000. *Objectivity and Insight*. Oxford: Clarendon Press.
- Sacks, O. 1985. *The Man Who Mistook His Wife for a Hat*. London: Duckworth.
- Sainsbury, R. M. 1979. *Russell*. London: Routledge.
- Sainsbury, R. M. 1985. Critical Notice of *The Varieties of Reference*. *Mind* 94, pp.120-142.
- Schaper, E. and Vossenkuhl, W. (eds.). 1989. *Reading Kant: New Perspectives on Transcendental Arguments and Critical Philosophy*. Oxford: Blackwell.
- Schopenhauer, A. 1966. *The World as Will and Representation*. Translated by E. F. J. Payne. 2 Vols. New York: Dover.
- Scott, M. 2001. Tactual Perception. *Australian Journal of Philosophy* 79, pp.149-16.
- Shoemaker, S. 1968. Self-Reference and Self-Awareness. *The Journal of Philosophy*, 65. Reprinted in Shoemaker, 1984, pp.6-18.
- Shoemaker, S. 1970. Persons and their Pasts. *American Philosophical Quarterly* 7. Reprinted in Shoemaker, 1984, pp.19-48.
- Shoemaker, S. 1984. *Identity, Cause, and Mind: Philosophical Essays*. Cambridge: Cambridge University Press.
- Shoemaker, S. 1986. Introspection and the self. In *Midwest Studies in Philosophy* 10. Reprinted in Shoemaker, 1996, pp.3-24.

- Shoemaker, S. 1994. The Royce Lectures: Self-Knowledge and “Inner Sense”. *Philosophy and Phenomenological Research* 54. Reprinted in Shoemaker, 1996, pp.201-268.
- Shoemaker, S. 1996. *The First-Person Perspective and Other Essays*. Cambridge: Cambridge University Press.
- Shoemaker, S. 1998. Causal and Metaphysical Necessity. *Pacific Philosophical Quarterly* 79, pp.59-77.
- Shoemaker, S. and Swinburne, R. 1984. *Personal Identity*. Oxford: Blackwell.
- Sosa, E. 1987. Subjects Among Other Things. In *Philosophical Perspectives 1:Metaphysics*. Reprinted in Rea, 1997, pp.63-89.
- Stanley, J. 1998. Persons and Their Properties. *The Philosophical Quarterly* 48, pp.159-175.
- Stern, R. 2000. *Transcendental Arguments and Scepticism: Answering the Question of Justification*. Oxford: Clarendon Press.
- Strawson, G. 1997. The Self. *Journal of Consciousness Studies* 4. Reprinted in Gallagher and Shear, 1999, pp.1-24.
- Strawson, G. 1999. Self, Body, and Experience. *Proceedings of the Aristotelian Society, Supplementary Volume 73*, pp.307-332.
- Strawson, G. 2003. Mental Ballistics or the Involuntariness of Spontaneity. *Proceedings of the Aristotelian Society* 103, pp.227-256.
- Strawson, P. F. 1959. *Individuals: An Essay in Descriptive Metaphysics*. London: Routledge.
- Strawson, P. F. 1966a. *The Bounds of Sense: An Essay on Kant’s Critique of Pure Reason*. London: Methuen.
- Strawson, P. F. 1966b. Self, Mind and Body. *Common Factor* 4. Reprinted in Strawson, 1974, pp.169-177.
- Strawson, P. F. Bennett on Kant’s Analytic. *The Philosophical Review* 77, pp.332-339.
- Strawson, P. F. 1974. *Freedom and Resentment and other essays*. London: Methuen.
- Stroud, B. 1968. Transcendental Arguments. *The Journal of Philosophy*, 65, pp.241-256.
- Sutton Morriss, P. 1982. Some Patterns of Identification and Otherness. *Journal of the British Society of Phenomenology* 13, pp. 216-225.

- Swinburne, R. 1984. *Personal Identity: The Dualist Theory*. In Shoemaker and Swinburne, 1984, pp.1-66.
- Swinburne, R. 1987. The Structure of the Soul. In Peacocke and Gillett, 1987, pp.33-55.
- Taylor, C. 1978-9. The Validity of Transcendental Arguments. *Proceedings of the Aristotelian Society* 79, pp.151-165.
- Taylor, C. 1979. Action as Expression. In Diamond and Teichman, 1979, pp.73-89.
- Teichmann, R. (ed.). 2000. *Logic, Cause and Action: Essays in Honour of Elizabeth Anscombe*. Cambridge: Cambridge University Press.
- Travis, C. 1994. On Constraints of Generality. *Proceedings of the Aristotelian Society* 94, pp.165-188.
- Van Straaten, Z. (ed.). 1980. *Philosophical Subjects: Essays Presented to P. F. Strawson*. Clarendon Press: Oxford.
- Velleman, D. 1996. Self to Self. *The Philosophical Review* 105, pp.39-76.
- Wettstein, H. 1991. *Has Semantics Rested on a Mistake? And Other Essays*. Stanford: Stanford University Press.
- White, A. R. 1968. *The Philosophy of Action*. Oxford: Oxford University Press.
- White, M. J. 1979. The First Person Pronoun: A Reply to Anscombe and Clarke. *Analysis* 39, pp.120-3.
- Wiggins, D. 1968. On Being in the Same Place at the Same Time. *The Philosophical Review* 77. Reprinted with 'Postscript' in Rea, 1997, pp.3-9.
- Williams, B. 1961. Strawson on Individuals. *Philosophy* 36. Revised and reprinted in Williams, 1973, pp.101-126.
- Williams, B. 1973. *Problems of the Self*. Cambridge: Cambridge University Press.
- Williamson, T. 2000. *Knowledge and its Limits*. Oxford: Oxford University Press.
- Wittgenstein, L. 1953. *Philosophical Investigations*. Translated by G. E. M. Anscombe. Oxford: Blackwell.
- Wittgenstein, L. 1958. *The Blue and Brown Books*. Oxford: Blackwell.
- Wright, C. 1992. *Truth and Objectivity*. Cambridge, Mass.: Harvard University Press.
- Wright, C. 1998. Self-Knowledge: The Wittgensteinian Legacy. In Wright, Smith and MacDonald, 1998, pp.13-45.

Wright, C., Smith, B. and MacDonald, C. (eds.). 1998. *Knowing Our Own Minds*. Oxford: Clarendon Press.

Zahavi, D. 2000. *Exploring the Self: Philosophical and Psychopathological Perspectives on Self-Experience*. Amsterdam: John Benjamins Publishing Company.