

Cartesian Self Representation

by

Eric Twa

A thesis submitted to the Faculty of Graduate and Postdoctoral Affairs in
partial fulfillment of the requirements for the degree of

Master of Arts

in

Philosophy

Carleton University

Ottawa, Ontario

© 2012

Eric Twa



Library and Archives
Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

ISBN: 978-0-494-93616-0

Our file Notre référence

ISBN: 978-0-494-93616-0

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Abstract

The thesis explores Cartesian self representation. Specifically, it distinguishes between two kinds of self representations: self referential representation and self presenting representation. Self referential representation is when there is a reference back to the I or ego. Self presenting representation is when there is a representation that also includes a representation of itself. I argue that all conscious representations include a self referential representation, while most (but not all) conscious representations include a self presenting representation. This distinction between self referential representations and self presenting representations is important, because many people confuse self presenting representations for self referential representations, arguing that just because self presenting representations are not present, no self representation is present either. At the end, I apply my thesis of self referential representations to the problem of how we gain knowledge of the external world. While I do not try to solve the problem itself, I show that self referential representations need to be a part of the solution.

Acknowledgements

I would like to thank Dr. Andrew Brook for all of his helpful insight.

I would like to dedicate this thesis to my Parents Pam and Allan Twa

Table of Contents

Abstract	2
Acknowledgements	3
Table of Contents	4
1. Introduction	5
2. A Brief Look at Descartes	12
2.1 Rethinking Descartes	13
2.2 Rejection of the Cartesian theatre	13
3. But why are self referential representations explicit?	18
3.1 Perry	18
3.2 Anscombe	24
3.3 Campbell	29
4. Incorrigibility	34
4.1 Armstrong	35
4.2 Costa	39
5. Self Presenting Representations	43
5.1 Self Presentation in Knowing that	44
5.1.1 Blindsight	52
5.1.2 Inattentional Blindness	57
5.2 Self Presentation in Action	60
5.2.1 The Dual Route Theory of Vision	63
5.3 Self representation	67
6. Our Knowledge of the external world	76
6.1 Ismael and Situated Representation	77
6.2 Campbell's Critique of Cartesian Representationalism	82
6.3 On Hallucinations	87
Bibliography	99

1. Introduction

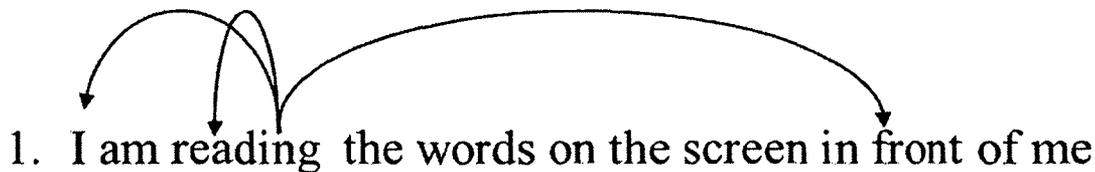
As I sit here viewing the screen there are representations of the screen and the words as they appear on the screen. And as I type these words there is a representation of my hands moving across the keyboard. But is that all there is? Specifically, is there only a representation of the screen, the words on the screen, and the keyboard? Or, are all of these representations actually representations of the world *in relation to me*? In other words, do I have a place in these representations of the world?

The traditional view of self representation is found in the Cartesian theory of the ego. In Descartes' statement, 'I think therefore I am,' Descartes argues that there is an explicit self-representation in consciousness and that this self representation is a part of all conscious representations. I am in favour of this claim. But, to be clear, I am not in favour of the entire Cartesian perspective. Descartes was a dualist and I am not supporting that thesis. Despite these differences, I will argue that Descartes' perspective of self representation is a valid perspective that requires further elaboration and defence from critics.

But what is meant by the term 'self representation'? Descartes clearly wanted it to mean that we are conscious of our own existence. But what is meant by 'our own existence'? In Descartes' terminology, he meant the existence of the cogito or I. He argues that there is a self and that this self is immediately apparent to us and that consciousness of its existence is incorrigible. But, as we shall see in Chapter 2, Descartes meant more than simply one's consciousness of one's own existence. He also meant that we had knowledge that we are having certain kinds of experiences and have certain cognitive abilities. For Descartes, the cogito was an active, thinking thing and not just an epiphenomenal I. Also, Descartes does not limit self representation to symbolic thought. Instead, he places self representation in all aspects of

conscious experience. But this makes it seem like there is more than one kind of self representation. If so, it will be useful to address each kind of self representation individually.

But is there more than one kind of self representation? I think that there is. Specifically, I think that there are self referential representations and representations that represent themselves (these are sometimes called self presenting representations). Perhaps this is best elaborated with an example. If we take the proposition, ‘I am reading the words on the screen in front of me’. There seems to be multiple kinds of representations. Specifically, there seem to be three representations. First, there is the representation of the words on the screen. There is the possibility of the representations of the I that is doing the reading. And there is the possibility of a representation of the reading itself. This can be visually depicted by a diagram provided to me by Andrew Brook.



In this diagram we clearly see that there are three kinds of representations. First, there are the words on the screen, which is not a kind of self representation. Rather, it is a representation of the world. Second, the arrow that points back to the “I” is a self referential representation, because it refers back to the self. Third, the arrow that points back to the “reading” is a self presenting representation, because it is the representation that represents the world and also represents itself. I will elaborate on each of these in greater detail.

The claim that there are representations of the world states that there are representations of the various objects in the world. As I sit here typing on the computer there are representations

of the words as they appear on the computer screen. And as I turn to grab my coffee cup, there is a representation of the cup. As I reach out my arm and grab the cup, there is a representation of the feeling of the cup. Then as I put the cups to my lips and take a sip, there is a representation of the coffee and the tastes of the coffee. The idea that there are representations of the world is not controversial. But is that all there is to representations? Some hold a theory called representationalism, which claims that this is all there is to representations. There is nothing more than the representation of the world. However, I want to argue that there are self referential representations and self presenting representations as a part of these representations of the world. The existence of self presenting representations is controversial. However, there is clearly a conceptual distinction between seeing the object and knowing that you are seeing the object. As this paper continues, I will give several examples for self presenting representations. But, generally speaking, I will be assuming their existence and focusing on the distinction between self presenting representations and self referential representations.

The thesis that all representations include self referential representations states that there is a reference to the I in our representation of the world. For example, as I sit here typing on the computer there is the representation of the words on the screen *in relation to me*. And this relation to me is an explicit (i.e. conscious) part of the representation of the words on the screen. As I turn to my coffee cup, the coffee cup is represented in the world in relation to me. As I reach out to grab the coffee cup I know how far to reach because of this referential representation of the coffee cup in relation to me. As I grab the cup there is a representation, and feel, of the coffee cup *in my hand* and my hand is a certain position in relation to the rest of my body. It is this relation that allows me to bring the cup up to my mouth and take a sip. As I take a sip, there

is the representation of the coffee in my mouth, so all of these representations of the world have a relation to me.

It's also important to note that self referential representations can exist without self presenting representations. We can have a conscious experience of an object without knowing that we are having a conscious experience of that object (without having a self presenting representation). But we cannot have a conscious representation of an object without having a point of view built into that conscious representation. So, while self presenting representations are not a part of *all* conscious experiences (although I will argue that they are a part of most conscious experiences) self referential representations are a part of all conscious experience of the world.

Since I consider self referential representations to be a part of *all* conscious representations, I will spend significantly greater time elaborating on self referential representations than self presenting representations. However, self presenting representations are also a kind of self representation and are a part of the traditional Cartesian thesis of self representations. So I will also be addressing self presenting representations.

Self presenting representations are representations that represent themselves. As I sit here typing on the computer, I have a representation of the words on the screen and I am aware that I have a representation of the words on the screen. I know that I am seeing the words on the screen. When I turn to my coffee mug, I have a representation of the coffee mug and I am aware that I have a representation of the coffee mug. As I reach out and grab the coffee mug I have a representation of my arm moving towards the coffee mug and I am aware of my arm moving towards the coffee mug. I know that I am reaching out to grab the coffee mug. And so on. But

self presenting representations can be broken up into two different kinds of representations. First, there is the self presenting representation of knowing that. Examples of this are the self presenting representation of the words on the screen or the coffee cup. They are representations where we are aware that we are experiencing a representation, because I both see the screen or the coffee cup and know that I am seeing the screen or coffee cup.

Second, there is the self presenting representation that I have a conscious ability. An example of a conscious ability is that I can distinguish between two colours. When I see red and blue, I know that I am seeing two colours and I know that I can distinguish between the two colours. In fact, I cannot help but distinguish between the two colours. But there are also the self presenting representations in action more generally. As I ride a bike, I have a representation that I am riding a bike and as I reach for my coffee cup there is a representation that I am reaching for my coffee cup. To be clear, we do not have a self presenting representation of every little detail that goes on in our action. I do not have a self presenting representation of all the specific movements in my hand while I am typing on the computer. But I do have a self presenting representation of the fact that I am typing. Or, as I ride my bike, I do not have a self presenting representation of every little detailed adjustment that I make as I ride. But I do have a self presenting representation of my riding a bike. As we will see in chapter 5, sections 5.2 and 5.2.1, the self presenting representations that take place in action are essential for controlling one's actions. This means that the self representations play an integral role in our interaction with the world.

The thesis will unfold as follows: in chapter 2, I will introduce Descartes' thesis of self representation. The point will not be to elaborate on Descartes' entire theory. Instead, I will focus on the self representation that is expressed in the statement, 'I think therefore I am'. And

elaborate on Descartes meaning of the word ‘think’. I will show that Descartes does not hold a Cartesian theatre view of consciousness, where the experiences of the world are separate from the self that is having the experience. Instead, he argues that the experiences of the world are a part of the I that thinks. This means that, for Descartes, self referential representations are not just a part of symbolic thought. Instead, they are a part of all aspects of conscious experience. Then I will outline Descartes’ theory of self referential representations and self presenting representations, showing that Descartes argues that both of these kinds of self representation are a part of all aspects of conscious experience. And not just symbolic thought.

In chapter 3, I will argue for the explicit (i.e., conscious) nature of these self referential representations against three critics. Anscombe argues that the “I” has no referent; Perry argues that the “I” has a referent but it is unconscious; and Campbell builds on both of these positions. By analyzing these positions; I will show where they have faults and, from this, why self referential representations are explicit.

In chapter 4, I will address the incorrigibility of self referential representation. Historically, many have argued that all mental states are incorrigible to the person experiencing them. Lately, however, there has been a backlash against incorrigibility. Some have even argued that no mental states are incorrigible. I take a middle position on this topic and argue that self referential representation (and only self referential representations) is incorrigible. Because of this, I limit my analysis to critique of the incorrigibility of self referential representation and do not address all arguments for and against incorrigibility.

In chapter 5, I include self presenting representations in my analysis of self referential representations. I believe that self presenting representations are a part of most conscious

representations. However, I do not believe that they are a part of all representations. In my analysis of self presenting representations I will show that, even when there are no self presenting representations, there are still self referential representations.

Chapter 6 consists of an application of self referential representations to the problem of how we gain knowledge of the external world. In this chapter, I will argue for a position of mild external world skepticism, while looking at arguments for perceptual realism (what Campbell calls the Relational View) and Ismael's theory from situated cognition. This application concludes my thesis.

2. A Brief Look at Descartes

This chapter will unfold in two parts. The first part will comprise a rethinking of Descartes' conclusions that we are "thinking things". Specifically, it will challenge the view of the Cartesian perspective as a Cartesian theatre, where the experiences of perception and the body are in relation to the ego but are not a part of the ego. After this rethinking there will also be an analysis of Descartes' thesis of self representation. If the Cartesian ego includes our experiences of perception and the body, then the self representation that is outlined in Descartes argument 'I think therefore I am' is the same in perception and the experience of the body as it is in the experience of symbolic thought. It is this part, the fact that there are self referential representations in all conscious representations and self presenting representations in most conscious representations, which I want to expand on, both in this chapter and in this thesis

To clarify the purpose of this chapter (and this essay), I am not arguing for an exact replica of the Cartesian ego as Descartes depicts it, nor am I trying to elaborate on Descartes' theory of the self. For example, I do not accept the thesis of substance dualism, where the ego is made of non physical spirit. Nor do I want to suggest that the ego is necessarily indivisible. The purpose of this chapter is to elaborate on the self representational aspect of Descartes' thesis, because I argue that this aspect of the ego Descartes got right. The reason for discussing Descartes' is that his statement 'I think therefore I am' includes self referential representations and self presenting representations. This makes Descartes a good launching pad into the main discussion of self representation, which is the main focus of this thesis. I could have discussed other historical philosophers instead. Or I could have discussed multiple historical philosophers. I have chosen to focus on Descartes and only Descartes because I believe that his statement, 'I think therefore I am', is the best historical statement for self referential representations.

2.1 Rethinking Descartes

Descartes develops his view of the I by means of a thought experiment. He comes to realize that he is sometimes deceived by his senses and that there are times when he believes that he is certain of something only later to realize that that belief is false. So he decided to doubt everything that can be doubted until he finds something that is certain or he finds out that there is nothing certain. He does this by proposing the possibility of an evil demon that is deceiving him in all of his endeavours. From this he concludes that there may not be an external world, because the world could be nothing more than a hallucination and that he may not actually have a body, because that too could be caused by a hallucination. In the same way that a person who has gone through an amputation can still experience phantom pains in the missing limb, he may be experiencing a phantom body and phantom world that does not actually exist. But he then comes to conclude that there is one thing that he is certain of, namely, that he exists. He is certain of this because the very act of doubting proves the existence of a thing that doubts. After all, one cannot doubt the existence of all things (even of oneself) if there is not something there to do the doubting. He calls this thing that doubts the I and further concludes that the I can also imagine and have sensory experience, because even if the external world does not exist, the experience of an external world still exists. In short, Descartes concludes that the I is a thing that thinks.

2.2 Rejection of the Cartesian theatre

Some interpret the Cartesian perspective in a way that is slightly different from the one above. In this interpretation the Cartesian ego is made like a theatre. This is called the Cartesian theatre theory of consciousness. In the Cartesian theatre the ego or I is sitting in a theatre. Her perceptual experiences are not a part of the I but displayed in front of the I, as if on a movie screen. The auditory experiences are also not a part of the I but are played to the I, as if from speakers. In short, the perceptual experiences are disassociated from the I and the experiences of

the body are basically an alien entity. Dennett (1991) is an example of the Cartesian theatre interpretation of Descartes. Specifically, Dennett has two criticisms of Descartes: First, Dennett critiques Descartes' dualist position. Second, Dennett critiques the subject/experience disunity that Dennett argues is a part of the Cartesian theatre. I have no objections to Dennett's critique of dualism, so I will focus on Dennett's second critique. Dennett introduces the second critique with the following statement:

You may well suspect that under cover of antidualism (“Let’s get that spook stuff out of here!”), I have spirited away (quite literally) something Descartes was actually right about: There is a functional place of some sort where the items of phenomenology are ... *projected* (1991, 126-27: his emphasis)

Dennett furthers his critique of Descartes by arguing that “*there is no such theatre, there is no such audience*” (1991, 128: his emphasis). From this we can clearly see that Dennett depicts Descartes' perspective of the self in the way that I describe it above. Where there is a self sitting in the theatre and our experiences are played to us.

However, a careful look at Descartes' own words shows that this is a misinterpretation of the Cartesian view. I will look at two quotes specifically:

But what then am I? A thing that thinks. What is that? A thing that doubts, understands, affirms, denies, is willing, is unwilling, and also imagines and has sensory perceptions (1941/1988: 83)

And in his elaboration on the I, he states:

Lastly, it is also the same ‘I’ who has sensory perceptions, or is aware of bodily things as it were through the senses. For example, I am now seeing light, hearing a noise, feeling

heat. But I am asleep, so all this is false. Yet I certainly *seem* to see, to hear, and to be warmed. This cannot be false; what is called 'having a sensory perception is strictly just this, and in this restricted sense of the term it is simply thinking (1641/1988: 83)

From these two quotes we can clearly see that Descartes includes the experiences of perception and the body in his definition of the "I" and "thinking". This means that Descartes does not use the usual definition of "thinking" that we are all used to. We usually consider thinking to be separate from seeing but Descartes does not.

To better demonstrate what Descartes is getting at, I propose we use the term "experience" instead of thinking. So 'I experience therefore I am' is a rewriting of Descartes' 'I think therefore I am'. This is important in the following respect. First, despite the belief of some, Descartes did not actually envision the I as a theatre. The experiences of the body and of the world are not disassociated from the I. This means that perceptual representations and bodily representations are a part of the ego and do not simply relate to the ego. And the self representations that are a part of every representation are also a part of the ego.

Also, Descartes' perspective does not limit itself to the self referential part of representations. As the quotations above demonstrate, Descartes position also entails that we have a knowing that and a knowing how (such as the ability to think and perceive). I will address knowing that first. Many philosophers argue that all (or at least most) conscious experiences are about something. This is usually called the intentionality of the mental state. And the intentionality of the mental state is usually defined as a knowing that. However, there are examples of conscious experiences that do not seem to have intentionality. Specifically, generalized anxiety disorder is a conscious experience but the experience of anxiety is not about

anything in particular. Despite this, someone could argue that an experience of generalized anxiety is simply about the state of anxiety itself. Whether or not we should accept this interpretation is questionable. But even if we reject this interpretation, we can still say that every conscious experience is a specific kind of experience. For example, a visual experience is always an experience of a certain colour or shade; it has certain dimensions and is in a certain relation to me. As Descartes points out, I may be tricked by an evil demon and the experience is a hallucination. But I still have a knowing that in regard to the experience itself. An objection to this could come from the experience of “nothingness” that is achieved in deep meditation. Unfortunately, I’ve never had such an experience so I’m not sure what the experience is like, but I assume that there is something that it is like to be in this experience. And if there is something that it is like to have the experience than we can know that this is what it is like to have the particular experience. If so, then this experience of “nothingness” is itself a kind of experience.

With regard to knowing how, most (and possibly all) of our conscious experiences are capable of performing a function. When I know what it is like to see a colour, then I know how to pick out that colour at a later date. I know how to imagine that colour. And distinguish that colour from other colours. Lewis (1988) even goes so far to argue that what it is like is nothing more than a knowing how. While I reject such a strong position, but there is good reason to believe that a knowing how is a part of our conscious experience. What’s important for our interpretation of Descartes is the acknowledgement that Descartes did not have a purely intellectualist conception of the mind. Instead, Descartes’ had an interactionist theory of mind, where the mind is an acting subject. Of course, Descartes was a dualist interactionist and we are not. But the fact that the mind is an acting mind will be important later on in the paper.

Knowing that and knowing how are closely linked, even though they are not one and the same. In every knowing how there are knowing that and every knowing that is a part of a knowing how. To elaborate, if I know that I am seeing two colours then I know how to distinguish between the two colours. And if I know how to distinguish between two colours then I know that there are two colours. We cannot have one without the other. For example, knowing how to ride a bike cannot be reduced to knowing that's. No matter how much book knowledge we have of bike riding, we still do not know how to ride a bike. But we know how to try and ride a bike. We know not to try and sit on the handle bars and steer with the peddles. In the same way, in knowing how to ride a bike we know several knowing that's. We know that you steer with handle bars and we know that you create momentum with the peddles. So knowing that and knowing how have significant overlap.

Also, there are other issues that need to be addressed. First, my reinterpretation of Descartes focuses on consciousness. There are unconscious parts of the ego and there is no incorrigibility in these unconscious representations. Second, my claim of incorrigibility is only related to the self referential representations in conscious representations. But self referential representation is not as self evident as Descartes made it out to seem. Some philosophers—in the next chapter I will focus on the work of Perry, Anscombe, and Campbell—have argued that there can be a conscious experience with either no reference at all or no conscious reference to the self and, therefore, no self referential representation. I will address this first before I address the incorrigibility of self representation because there is no point in arguing for its incorrigibility if it doesn't exist.

3. But why are self referential representations explicit?

Is there a self referential aspect to representations? Some have argued that there isn't actually a self representational aspect to representations but that, instead, we acquire knowledge of the world without any representation to the self. This is the position of John Perry (1986, 1997) and John Campbell (2004). And G.E.M. Anscombe (1975) argues that the "I" has no referent. I will address Perry first.

3.1 Perry

Perry acknowledges that in our perceptual experiences of the world there seems to be a relation *to me*. For example, as the reader reads this page, the page is in a certain position in relation to him or herself. He or she sees the page from a certain angle, it is at a certain distance, and sitting on the table or held in ones hand in a certain way. However, Perry expresses sympathy with Wittgenstein and Lichtenberg's argument that there is no concept of the self in these experiences, and that we are better off to say that 'it thinks' instead of 'I think'.

Specifically, Moore describes Wittgenstein's position as follows:

...the point on which he seemed most anxious to insist was that what we call 'having toothache' is what he called 'a primary experience' ...; and he said that what characterizes 'primary experience' is that in its case, "I" does not denote a possessor'. In order to make clear what he meant by this he compared 'I have a toothache' with 'I see a red patch'; and said of what he called 'visual sensations' generally ... that 'the idea of a person doesn't enter into the description of it, just as a (physical) eye doesn't enter into the description of what is seen'; and he said that similarly 'the idea of a person' doesn't enter into the description of 'having toothache' he said that 'just as no (physical) eye is involved in seeing, so no Ego is involved in thinking or having toothache'; and he

quoted, with apparent approval, Lichtenberg's saying, 'Instead of "I think" we ought to say "It thinks"... (Moore 1962, 302-3 cited in Perry 1986 137-8)

The goal of Perry's paper "Thought without Representation" (1986) is to try and explain how this is possible. I hope to demonstrate that it is not possible for references of the self. I will conclude that "I think" is the proper description.

The basic idea expressed in Perry's 1986 essay is that there are unarticulated constituents in thoughts and propositions. Perry starts with an example, 'It is raining'. To elaborate, if I ask you 'how's the weather?' and you say, "It is raining" then I know that you mean, "It is raining *here*". But if you are on the phone with a friend who lives somewhere else and I ask you how the weather is, you can respond with the same statement, "It is raining", but you can mean "it is raining *there*". Now, in order to establish the truth value of these propositions, we have to analyze the constituents. And the unarticulated indexical "here" and "there" are essential for determining the truth value for the proposition "it is raining".

To explain his position, Perry introduces an example of Z-Landers. Z-Landers live in only one location, they do not travel, and they do not talk to anyone outside of their location. So, for Z-Landers, there is no need for indexicals like "here" or "there". When a Z-Lander says that "it is raining" everyone in Z-Land automatically knows to grab an umbrella when they are going out or to stay inside. Further, the Z-Landers could slowly migrate west at such a pace that they do not notice that they are migrating. Still, the proposition that "it is raining" will migrate with the Z-Landers as they move west. So, for the Z-Landers the indexical "here" is an unarticulated constituent that, because of their isolation, never needs to be expressed.

At first glance, it looks like unarticulated constituents are simply context sensitive. This would mean that it is the context in which the utterance is made that supplies the constituent. However, Perry thinks up an example of some chauvinistic San Franciscans that have made up their own particular dialect. In this dialect the statement, “it is raining”, always refers to San Francisco—no matter where it is uttered. In this particular example, there is still an unarticulated constituent, because “it is raining” means that “it is raining *in San Francisco*”, but it is context insensitive. So the theory of an unarticulated constituent cannot always be explained by context sensitivity. Instead, in this situation the unarticulated constituent is explained by the structure of the language itself.

But is the referent of “I” (and the self representation that goes along with it) explained by the structure of the language or is it context sensitive? It is clearly context sensitive, which makes it an indexical. If it was explained by the structure of the language then the referent would not change from context to context. But “I” always refers to the person speaking. So it does change from context to context. The fact that the “I” is an indexical will come up again later in this essay.

Unarticulated constituents are the basic principle behind Perry’s theory that there can be thoughts without representations. And he applies the idea to conscious representations. Basically, Perry argues that there is a distinction between being aware of X (be it a perceptual experience or an internal emotion) and being aware of the subject that is aware of X. In the first instance, the self representational aspect is like an unarticulated constituent. In the second instance, the self representation is articulated. I want to argue that this is only a conceptual difference, not a difference that is found in the real world. To do this I need to demonstrate that the self

representation that is involved in conscious representations is more explicit than the unarticulated constituents in statements like, “it is raining”.

First, I’d like to point out that, when it comes to unarticulated constituents, we can make mistakes. If someone tells me that “it is raining” I can accidentally assume that they are talking about somewhere else when they are talking about here, or I can accidentally assume that they are talking about here when they are talking about somewhere else. However, when it comes to my current conscious experience—for an example, I will focus on visual experience—there is no doubt that the experience is my own. I will not mistake it for someone else’s. This seems to indicate that the self representation is more explicit than the unarticulated constituents that Perry talks about. After all, if they were unarticulated constituents that are context sensitive, then, like all other unarticulated contexts constituents that are context sensitive, there should be a possibility for confusion and error. When someone else states that “I see,” I should be able to think that they are talking about a third person. But as long as she knows the proper use of the term “I,” this cannot happen.

Of course, it is possible, at least in theory, for me to “recall” someone else’s memory. For example if someone were to transplant a part of someone else’s brain into my brain while I was sleeping. And that part of the transplanted tissue stored memory. I could wake up and “recall” participating in an event that I did not actually participate in. However, I am arguing that it is not our memories that have an explicit self referential representation. It is our current conscious experiences. Not only can I have someone else’s memories transplanted into my brain, I could have false memories and distorted memories as well. But when I recall the memory, even if it is a false memory or some else’s memory, *while I am experiencing the memory*, is clearly my own experience and not someone else’s.

The second problem follows from the first problem. In unarticulated constituents there are several different constituents that can be used. The statement, “it’s raining”, can have the unarticulated constituent of “here”, “there”, “Calgary”, or any other place. Similarly, when I say that it is 2:00 PM, it usually includes a presupposition of the time zone that we are in. But I could be referring to Mountain Time, Pacific Time, or any other time zone. But when it comes to the self referential representations of conscious experience, the referent is always the same. It doesn’t change. This means that self referential representation cannot work like demonstratives or context sensitive demonstratives. But they also do not work like Perry’s example of unarticulated constituents that are context insensitive, such as his example of the San Franciscans, because the referent for the San Franciscans doesn’t change when the speaker changes. But the referent of “I” does change when the speaker changes.

A third problem arises in Perry (1997). While trying to elaborate why self referential representations are not explicit, Perry gives three examples. First, a child looks at a clock and sees that its 3 PM. This is true, given that it is 3 PM Pacific Time. But the child does not know that it is Pacific Time, it’s unarticulated, but the child is right in believing that it is three PM. Second, long ago Europeans had no concept of Northern Hemisphere and Southern Hemisphere and they wrongly believed that it was necessarily a part of the summer months that it is summer. But there was an unarticulated constituent in this knowledge. June, July, and August are summer months only in the Northern Hemisphere. But still, these ancient Europeans held a true belief. Third, there can be a case where a child does not distinguish between ‘here’ and ‘there’ and mistakenly believes that when it rains, it rains everywhere. But, despite the fact that the child has no knowledge of the unarticulated constituent that she expresses when she says, “It’s raining,” she still has true knowledge about the fact that it’s raining. Perry puts self referential

representations in the same boat. They can be a part of the representation in an unarticulated way, but are not explicitly a part of the conscious representation.

However, self referential representations are fundamentally different from the three examples above. It's true that a child can be right about the time and know nothing about time zones, ancient Europeans can be right about the summer months in Europe but know nothing about the hemispheres, and a child can be right about the fact that it's raining without realizing that it's only raining here. But people—even a small child—cannot have a visual experience without knowing about their own existence and their experience in relation to the external object, since the child always sees the world from a particular point of view. And this point of view dictates what we can and cannot see. It dictates what is to the left of us, what is to the right of us, what is above eye level, and what is below eye level. So it sets the frame of seeing, it constrains what the child can see (he or she knows that he or she needs to turn her head in order to see stuff that is outside of her point of view). The self referential representation is a conscious part of the experience of the world. The child actively experiences the point of view as a part of the experience. Further, a small child would not mistakenly believe that, because an apple is a meter in front of her, it is a meter in front of everything, even though the TV is, let's say, five meters in front of her. It is because self referential representations are explicit that these mistakes do not happen. And the self referential part of our conscious experience allows us to have a concept of distance in relation to us in our visual experience. She knows that the TV is five meters away from her and that the apple is only a meter away from her. If there was no self referential representation (or it was unconscious) then these distances would not be a part of the conscious experience.

When we have a point of view (I will focus on vision) and the point of view plays a role in our experience of the world, then the point of view contains an explicit (i.e. conscious) self referential representation. In every visual experience we see the world from a specific angle. Further, we do not see everything. Instead, our point of view frames what we can and cannot see. If we want to see something that is outside of this frame, then we have to turn our head or move our body in order to change our point of view. If we want to focus on something that is off to the side, but still in our visual experience, then we have to turn our eyes or our heads. Since our point of view sets the frame that we see, it is an explicit part of our conscious experience of the world. We necessarily see objects as directly in front of us, to the left of us, or to the right of us. We might not know the words “left” and “right” but the fact that the objects have a particular directional relationship to us is explicitly a part of the representational experience of the objects in the world. Without this knowledge we would not be able to act on the objects. So evidence that we can act on the objects is evidence that a point of view is playing a role in the action. This is not learned either, we instinctually know that we have a point of view and that it plays a role in our seeing the world. People have to learn about how the indexical ‘here’ and ‘there’ work, about the difference between the seasons and the northern and southern hemisphere, and about time zones. These are unarticulated constituents of thought. But people do not have to learn about their own point of view and about the fact that their own point of view is different from other people’s point of view. This self representational aspect of consciousness cannot be an unarticulated constituent. Instead, it must be an explicit part of our representations of the world.

3.2 Anscombe

In “The First Person” (1975), G.E.M Anscombe argues that the “I” is an empty term, meaning that it does not have a referent. Specifically, she argues that there is no self-consciousness. By self-consciousness, Anscombe means what I have been calling self referential

representation. To be clear, the word 'I' has a use, just not a referent. Traditionally, the "I" is thought to refer to the self of the person speaking. This means, first, that there needs to be a self that can be referred to and, second, that we need to be conscious of the self. Anscombe calls this self consciousness, but I have been referring to it as self referential representations. First, I will briefly outline Anscombe's argument for why "I" is a term without a referent. Then I will work backwards and try to uncover where Anscombe went wrong in her analysis of the first person. I will not be addressing Anscombe's argument in its entirety, because the paper makes several points that are irrelevant to the discussion at hand. Instead, I will be focusing on the relevant claims for the discussion. And leave out other claims that are either irrelevant to the discussion or are claims that I agree with.

To make this argument she analyzes several possible ways that "I" could have a referent, on some of them I agree with. For example, she argues that "I" cannot be a proper name. Not simply because it is used by everyone (in theory, everyone could have the same name. It would be very confusing, but there's no law against it), but, because I cannot use "I" to refer to you and you cannot use "I" to refer to me. If "I" was a proper name than I should be able to use it to refer to you in the same way that I can use your other names (even nick names) to refer to you. She also argues that "I" cannot be a demonstrative because other demonstratives, like "this" or "that" are empty concepts when they are taken on their own (I will criticize Anscombe's very short analysis of indexicals later in this section). These analyses (along with a few others that are irrelevant to this critique) lead Anscombe to conclude that "I" is an empty term that has no referent and that there is no such thing as self consciousness. This causes her to conclude that statements such as, "I am Anscombe" or "I am Eric Twa" are not actually identity statements.

I will start my critique with the conclusion and work my way backwards. If “I” is an empty term that has no referent then we have a problem explaining how certain beliefs work. Specifically, in “The Problem of the Essential Indexical” (1979), John Perry points out that the “I” plays a specific role in our belief states. He does this through an example. One day he was in the super market and he sees a trail of sugar on the ground. He starts to follow the trail of sugar to try and find the person who is making the mess. As he goes in circles the trail of sugar gets bigger and bigger, until he stops and realizes that “I am making the mess” or “I am the man who is making the mess”. It is then, and only then, that he stops his cart and adjusts the bag of sugar so that he is no longer making a mess.

It is important to elaborate on this example to point out exactly what the “I” is doing. When he is thinking that someone is making a mess, he has a true belief. But this does not cause him to stop and fix the bag in his cart. If he sees himself in the mirror, but doesn’t recognize himself, he can point at the person in the mirror and say, “that man is making a mess.” Once again, he has a true belief. But this doesn’t stop him from fixing the bag of sugar in his cart. Even if he thinks “John Perry is making a mess” this still will not cause him to stop and fix the bag in his cart *unless* he also holds the belief that “I am John Perry”. So the belief that “I am John Perry” or “I am Eric Twa” needs to be an identity statement, because it plays that role in our belief system and behaviour. If the statement, “I am Eric Twa”, had no content, then we are at a loss to explain how this works.

The problem is Anscombe’s analysis of indexicals. She only analyzes demonstratives and not pure indexicals. Although, to be fair to Anscombe, the distinction between pure indexicals and demonstratives was made by Kaplan (1989), after she wrote the essay we are examining.

However, this new distinction that Kaplan has made causes real problems for Anscombe's analysis. Specifically, she states:

It used to be thought that a singular demonstrative, "this" or "that", if used correctly, could not lack a referent. But this is not so, as comes out if we consider the requirement for an answer to "this what?" Someone comes with a box and says "This is all that is left of poor Jones." The answer to "this what?" is "this parcel of ashes"; but unknown to the speaker the box is empty. What "this" has to have, if used correctly, is something that it *latches on to* (1975: 54).

From this Anscombe points out that calling the word "I" a demonstrative does not demonstrate that it has a reference.

However, since Kaplan (1989) we have a different view of indexicals and demonstratives. Words like "this" and "that" are true demonstratives, while words like "I" and "now" are not true demonstratives, rather, they are pure indexicals. Anscombe is correct to point out that a true demonstrative might not have a reference. We may be mistaken in our use of "this" and "that", thinking that they have a referent when they actually have no referent at all. But pure indexicals work in a different manner. Pure indexicals have a *fixed* reference. The word "now" is fixed by the time of the speaker. Whenever a speaker uses the word "now" they are necessarily referring to the time that they make the utterance. The word "I" is fixed by the person speaking. It always refers to the person speaking or the person thinking. So this argument that the "I" does not have a reference can be disproven by Kaplan's distinction between true demonstratives and pure indexicals. True demonstratives do not always have a referent, but pure indexicals *always* have a referent.

With that said, I want to go further into Anscombe's analysis of the first person to help elaborate how she got self referential representation wrong. Anscombe (1975: 49-50) gives a previous example of a word that seems to work just like "I", but, she claims, does not refer to the first person self. Imagine a society where everyone has a name written on the back of their shirt and high up on their chest so that they cannot see it. The names range from "B" to "Z". Also, everyone has "A" written on their wrist. Through socialization, everyone has learned to respond to the name that is written on their shirt. And when a person in this society wants to refer to someone else she uses the name written on their shirt. However, whenever a person wants to refer to themselves they use the name, "A", that is written on their wrist. In short, everyone uses the name "A" in the exact same way that we use the word "I" and everyone uses their other names, "B" to "Z", in the exact same way that we use proper names. However, despite this, Anscombe argues that the person's use of "A" in this society does not have the self consciousness that we want "I" users to have.

Anscombe acknowledges that the users of "A" have some type of self consciousness, insofar as they see their own actions, describe their own actions to others, and use "A" to refer to themselves. But she states that "when we speak of self-consciousness we don't mean that. We mean something manifested by the use of 'I' as opposed to 'A'" (1975: 50). But the only specific example she gives that would lead us to indicate that "I" is not self referential is that we could see one persons arm, and the "A" on their wrist, and mistakenly believe that it is our own. But the same is true about how my roommate could type out a note stating, "I need to get more milk" and I could look at it and wonder, "Did I write that?" The trick, and I do believe that it is nothing more than a trick, is that Anscombe has put the "A" on peoples' arms and that this allows them mistakenly believe that someone else's arm is their own. But we should note that if we all wrote

“I” on a wrist then we would be in the exact same situation. And we should note that when the people in this society use the word “A” in thought and in speech, it works just like the pure indexical “I”. It always refers to the speaker or thinker. So the people who use “A” instead of “I” and have “A” written on their wrists are actually self conscious and the “A” and the “I” has a conscious referent.

Anscombe tries to claim that self referential representations (what she calls self consciousness) are something that “I” users are supposed to have, but “A” users do not have. She then analyzes several possible ways to define the referent of “I”, concluding that there is no definition available. Before Kaplan’s distinction between pure indexicals and demonstratives this looked true. However, after this distinction, we see that “I” does have a referent and that there is no difference between “I” users and “A” users. So “A” users also have a referent.

3.3 Campbell

Much of John Campbell’s (2004) argument draws on Perry and Anscombe. Since I have already addressed both, I will not spend a large amount of time on Campbell. However, Campbell adds a few arguments that are worth addressing in detail. Specifically, he discusses relations such as ‘to the left of me’ and ‘to the right of me’ and he discusses the comparison to film. I will address the relations argument first.

Campbell analyzes the relations between the first person and the spatial content of perception and action. For example, we might say that the computer is in front of me. The dog is in front of the car. Or that the door is on Bill’s left. These relational notions specify who we are speaking about and the relation between that person and the object that we are speaking about. They are two-place notions: x is to the right of y , x is in front of y , and so on. So far so good, but it is here where Campbell takes a turn that I disagree with. He argues that when it comes to

describing our spatial vision, we do not need to include the fact that something is in front of or to the left of a particular subject. He states that certain animals may not be able to represent anybody else's right or left except their own. So the animal doesn't have an egocentric object that is always in its visual field. It doesn't have a representation of itself in relation to other objects. Instead, it's just represented as 'in front' or 'to the left'. These can be described in monadic egocentric terms such as 'x is to the left', 'x is below', and so on. Campbell describes these as "viewpointed" representations but states that "the representations themselves do not make it explicit whose viewpoint is being represented" (2004: 207). Then he states that this is the same for humans and this leads him to a familiar conclusion, namely, that we should claim that 'there is thought' instead of 'I am thinking'.

This leads Campbell to question why we use the self-referring term "I" at all. He suggests that we compare it to a logical constant such as 'and'. The logical constraint has elimination and introduction rules and, because of this, we know how to use the term. But there is more than simply these rules. The knowledge of the meaning of 'and' is provided by your grasp of the truth-tables for the term which validates the use of the rules. This view would align with Anscombe's view that there is no referent for the "I". However, I have already addressed Perry's and Anscombe's arguments and found them to be unsustainable. I will now respond to Campbell in greater detail.

But this is not how human interaction works. We do not simply give monadic relational descriptions such as 'to the left' or 'in front'. We continually describe things in two-place notation. Such as, 'x is to the right of y' or 'x is in front of y'. The fact that we do this so often and with such ease seems to suggest that this is a part of our perceptual experience of the world. Perhaps there are occasions where neurological damage has been done and a person cannot do this, but

when I see that my computer is one foot in front of me and I see that the wall is two feet in front of me, then I see that the computer is one foot in front of the wall. These two-place notations are an essential part of our understanding of the world and our interaction with the world. However, it's possible Campbell is only trying to explain how things relate to us in our own conscious experience. If this is the case, then we may not need two-place notions to describe the world. However, our position, and the visual point of view that we have, anchors us to a particular place. It's what makes something "to the left" or "to the right". This place is essential for how we see the world and cannot be removed from the experience of the world. So the self enters into our perceptual experiences of the world.

This means that the term "I" cannot be viewed in the same way as logical constraints such as "and". But this can also be demonstrated without the use of the conclusions above. There are no formal rules, such as the elimination and introduction rules for 'and', that can create universal truth tables for "I". This is because, as shown in my discussion of Anscombe, that the "I" is an essential indexical (meaning that it is a context sensitive term that always refers to the person speaking or writing). The context sensitivity indicates that there cannot be truth tables for the term in the same way that there are truth tables for conjunctives, disjunctives, conditionals, and bi-conditionals. So the "I" cannot be understood by its functional use and logical rules alone. In order to understand what that "I" refers to, we need to have a conscious understanding of its referent.

Campbell gives another argument in his paper that is worth addressing. To help elaborate his position, he gives the example of a "film" subject. I do not like this analogy and in chapter 2.2 on Descartes I reject the idea the consciousness or the ego should be viewed as a theatre. But to help my position, I will elaborate on Campbell's analogy and respond to it. Vision does not

just give us knowledge of the world it gives us knowledge from a certain position. I see the computer in front of me, the television diagonally to the right, the wall to the left, and so forth. And I do not need to continuously remind myself that it is my visual experience and that it is from my position. From this Campbell notes, "It is, though, a familiar point that in watching a film, you need not be constantly asking, 'and who is it that is observing all this?'" (2004: 210). And he applies this to documentary films. If a documentary is well done, you do not notice the camera and camera man. You see political debates, interviews, and police scuffle with rioters and you never think about how the documentary was made (you never think about the camera man and the camera). Indeed, Campbell notes that you can watch a documentary about two brave mountain climbers and believe that they are the only people on the mountain, you never even think about the camera man. He or she is an unarticulated constituent in the movie.

However, there are several problems with this analogy. It is true that you can watch a documentary of two mountain climbers without realizing that there is a camera man up there with them. But it is not true that people sit in a room with two other people and think that there are only two people in the room. People do not forget about themselves in the same way that they forget about the camera man. You would never climb a mountain with two other people, come back down, and accidentally report that only two people climbed the mountain. And the camera man himself or herself would never forget that he or she is actually up there on the mountain with the two climbers. We the viewers, from our detached perspective, may forget about the camera man but the camera man doesn't forget.

Basically, what Campbell tries to do is demonstrate that our knowledge of the referent of "I" is not like our knowledge of the referent of demonstratives like "that chair." To know what a demonstrative refers to you need to be conscious of the demonstrative itself. To know what "that

chair” refers to, you need to be conscious of the chair. But Campbell argues that this is not the case for “I”. Instead, we need to know how it is functionally used and that’s it. But my analysis of Perry and Anscombe show that this cannot be the case. The word “I” does not work like other unarticulated constituents. It is explicit. And the “I” does have a referent. In fact, its referent is more certain than that of a pure demonstrative. The demonstrative “that” can be hollow because a person can think that she is referring to something, but there is really nothing there. But because the “I” always refers to the speaker or thinker, and there cannot be a spoken or thought expression without there being a speaker or thinker, “I” always has a referent. And that referent is always a part of our conscious experience.

4. In corrigibility

Since the nineteen sixties the idea that our beliefs about our own representations are incorrigible has been widely criticized. I will not talk about all the different critiques of incorrigibility. Instead, I will focus on the critiques of Armstrong (1963) and Costa (2001) since their critiques directly apply to the Cartesian thesis of the incorrigibility of self referential representation. There are two reasons for this discussion on incorrigibility. First, I want to discuss all the different aspects of self referential representations laid out by Descartes and I consider incorrigibility to be a part of this discussion. Second, because incorrigibility has become widely criticized in recent years. We used to think that *all* mental states are incorrigible. Now many people think that no mental states are incorrigible. Both positions go too far. Most mental states are not incorrigible, but, I argue, self referential representation is an exception to this. To be clear, I am only arguing for the incorrigibility of one kind of self representation, the self referential aspect of conscious representations, and not for the incorrigibility of all mental states.

. By 'incorrigible' I mean that when the subject is conscious of her self referential representations she cannot be wrong about them. Self referential representations are incorrigible because when a person utters the statement "I exist" or "I think" he or she cannot be wrong. The very act of uttering or thinking the words demonstrates that there is a referent that confers the statement or thought. Further, if a person utters (or thinks) the statement "I do not exist" or "I do not think" he or she must be wrong. Since the very act of the denial disconfirms the statement or thought. As I noted at the end of section 3.3, when the "I" is used properly it *always* has a referent because it is impossible for there to be speech or thought without a speaker or thinker. And, since I elaborated on Descartes' position in chapter 2, demonstrating that by "think" Descartes was referring to all aspects of conscious experience we can expand these claims to the

rest of consciousness. To say that, “I see therefore I am” or “I feel therefore I am” is just as true as I think therefore I am.

But before I start on Armstrong’s critique, I need to address a critique about self deluded zombies. Some could argue that a deluded zombie has a “self representation” but it is not a genuine self representation because the zombie isn’t genuinely conscious. I do not want to digress too far into the zombie issue, because addressing the problem of zombies in full would take another essay. But I would respond that if a “zombie” really had a self representation, then it is not a genuine zombie. This leads to the problem about whether or not zombies are actually possible (I don’t think that they are). But I will leave this to another essay and simply sidestep the question.

4.1 Armstrong

Armstrong (1963) gives several arguments against the incorrigibility of representations. The first argument I want to address is Armstrong’s distinction between apprehension and the mental state being apprehended. Armstrong argues that this difference is necessary and because of it we can apply Hume’s argument of “distinct existences”. That if two things are separate then one can exist without the other. This means that we can believe that we have apprehended a mental state but the mental state is not actually there. Armstrong argues this position because “the apprehension of something must be distinct from the thing apprehended. For if not, we are faced with a flagrant circularity” (1963: 422). So the incorrigibility thesis is false. However, this draws on the fallacy of the “Cartesian theatre” that I address above. Descartes does not distinguish between the self and the experiences of the self. The experiences are a part of the self. However, Descartes does distinguish between the self and the world/physical body. So a person can be wrong about his or her visual experiences of the world or body, because these are

the things being apprehended, but when it comes to the experience itself, this is the apprehension. So to say that we can be wrong about our knowledge of the apprehension itself does not fall a foul of Hume's distinct existences.

Here we have the misinterpretation of Descartes' meaning of the words "I think". Thinking and the experience of thinking is a part of the self. If a man is thinking about an imaginary dog, there is a relation between the thinking and the imagined object. But the representation does not stand between the self and the object. There is not a relation between the self, the thinking, and then the imagined dog. Instead, the thinking is an action and an experience of the self in a certain state, the state of thinking, and this self is in a relation to the imagined object, the dog. In a similar fashion, when I see a red table, the representation does not stand between the self and the table. This is the view of the widely rejected sense-datum theory. But I do not argue for a sense-datum theory of perception. Instead, I argue that the experience of the red table is a part of the self in a certain state, the state of seeing a red table, and is in a relation with the actual table.

However, despite my rethinking Descartes, I could be wrong. I need to demonstrate that decouplability is not a universal criterion for representation. In short, I need an example where a decoupling cannot happen and then I need to demonstrate that self referential representations are the same. Mark Rowlands (2009, 126) gives an example from Andy Clark (1997, 144-145) about rats. Specifically, there is a population of neurons whose function is to carry information about the direction in which a rat's head is facing. Further, these neurons only function in tandem with a continuous stream of proprioceptive information from the rat's body and do not seem to satisfy the decouplability requirement. The self referential aspect of conscious representations works in a similar way. There is no self referential representation without representation and there is no

representation without self referential representation. For example, a visual experience *must* have a point of view. So the visual experience must include a self referential aspect to it. Note that I am not arguing for the existence of two different representations. Instead, I am arguing that representations of the world include a self referential representation as a part of the representation.

What is incorrigible is the self referential aspect of consciousness. Belief that the “I” exists is incorrigible. This knowledge of existence is what cannot be decoupled from the experience itself, because if there is an experience, the experience itself is the person having the experience. And there cannot be an experience without the existence of the experience. If there is no existence of the experience, then there is not experience at all. Even in a mistaken experience, such as a person thinking that they see red but they are actually seeing orange, there is still the existence of the experience and it is this knowledge of the existence of the experience that cannot be wrong.

A second objection from Armstrong that I would like to address goes as follows: suppose there is a brain technician who has complete knowledge of the workings of the brain. A person sincerely reports that she is in pain and the brain technician can use her belief to confirm that (1) she is not lying, (2) her brain is in a state for a different experience, (3) and there are disturbances in the brain process that account for the introspective mistake. I’m not sure if this is possible. But I will give Armstrong the benefit of the doubt and assume it is possible. Even if so, this argument would not raise a question about the incorrigibility of self referential representation. In Chapter 3, I demonstrate that all conscious representations have a self referential representation, so self referential representation is still a part of the mistake, because the person must have a hallucinatory experience that makes them believe that they are having an

experience that is not there. And the hallucinatory experience has a self referential representational aspect to it. This can be shown by the fact that we can act on our experiences, even our hallucinatory experiences. If you have a hallucination of a floating dagger you can reach out and try and grab that dagger. Your hand will go right through the dagger, but there is still a point of view between you and the hallucinated dagger. In other words, there is still a self referential aspect to the hallucination. The only way there isn't a self referential aspect is if there is no experience at all. But if there's no experience then there is no misrepresentation.

The third objection that I would like to address pertains directly to Descartes. It will be useful to quote Armstrong at length:

When Descartes set out to examine his thoughts, to see if any were indubitable [or incorrigible], he *presupposed* that he had these thoughts, for he could not examine his thoughts unless he had some to examine! But this did not make his statement "I am thinking now" a logically indubitable one. It was simply a logical presupposition of his starting point. In the same way, if it is *given* that I make a sincere report "I am in pain now," then that *presupposes* that I believe that I am in pain now. But that does not make the statement "I believe that I am in pain now" an indubitable one. It is simply a logical presupposition of the given starting point: a statement of what is in fact the case (1963: 426-427).

But Armstrong seems to misinterpret Descartes' point. It's not just that you can choose to presuppose your own existence before you start thinking but that you *must* know about your own existence when you are thinking. After all, thinking is an action and there must be something (in this case the I) that is doing the action in order for the action to take place. Perhaps this can be

better addressed with a specific example. A woman says, “I think therefore I am” and a brain scientist responds with, “No, you don’t actually exist, you just think that you do” or “No, you’re not actually thinking, you just think that you’re thinking”. This is just plain nonsense. And the same can be said for the self referential representational aspect in the experience of pain. In order to sincerely state that you are in pain, there must be a self referential representation that justifies the statement. The apprehension may be wrong. A woman can experience a pain in her foot even though there is nothing wrong with her foot in the same way that a war vet can experience phantom pains even though there is no limb at all. But the self referential representational aspect of the experience itself, and her belief about it, is incorrigible.

4.2 Costa

Costa (2001) focuses his critique on the incorrigibility of “I think” and then expands his critique to “I exist” as well. His argument unfolds via a number of claims.

Costa’s first argument is to point out that one always has to think about something. When we think, we are always thinking about something in the same way that when we wish or doubt we are always wishing or doubting something. Costa argues that it doesn’t make sense to say that ‘I’m thinking about thinking’ in the same way that it doesn’t make sense to say that I’m ‘wishing about wishing’ or ‘I doubt about doubting’. Of course, a person can think about their thinking processes or think about how they think, in the same way that a person can doubt their doubting processes or doubt one of their doubts. But this is still thinking about something. In this case, it’s the thinking about the act of thinking or about what a person is thinking about. No one is thinking while, at the same time, thinking about nothing at all. In common language, a person might say that they are thinking about nothing. But this is usually short for, ‘I’m thinking about nothing important’. When a person thinks they are always thinking about something.

Costa also notes that a person always thinks a whole thought. When we think we do not usually think thoughts that are not propositional. Further, for the thought to be meaningful the person must think a thought that's propositional. Costa argues that if the person's thought are not propositional then it isn't a true thought. It isn't a claim about the person or the world. Of course, we can have half a thought and stop ourselves halfway. But the act of thinking includes a whole thought. It's just that the whole thought isn't expressed, because the person changed their mind about the validity of the thought. This means that a thought is not instantaneous.

Costa then turns to his critique of the incorrigibility of "I think." The main crux of his thesis is that "*our consciousness is incapable of entertaining or focusing on more than one cognitive act at the same time*" (2001: 226, his italics). From this Costa argues that you cannot think about your thinking while also thinking about something else. To elaborate, Costa claims that when you think about something, you have to finish that thought before you can think about the fact that you are thinking. So he argues that since we can only think about one thing at a time, it isn't possible for us to think about the fact that we are thinking without finishing our thought and then reflect on the thought that we are having. This means that, for Costa, all self representation is memory and to have a self representational thought of any kind is to reflect on the memory of our experiences. But it is possible for us to be mistaken about what we are remembering. Since, if it's possible for us to forget about what we are thinking, then it's also possible for us to have a false memory that we were thinking when we were not thinking. And its possible for us to think that we existed even though we only have a false memory that we existed.

The first and most glaring problem with this is the idea that a mental representation cannot represent more than one thing at a time. According to this position, a mental

representation cannot represent the object of the thought and the act of thinking itself. But, if this were true, then we would not know that we are thinking while we are consciously thinking. Instead, we would have to wait until we finished thinking about a topic then reflect on our memory to learn that we are thinking. This seems implausible in two respects. First, there is no reason to believe that a representation cannot represent more than one object at a time. There is no reason to suggest that a representation cannot represent both an object and the representation itself. As I look at the computer screen, I have a visual experience of the computer screen and I know that I have a visual experience of the computer screen. There is a representation of the computer screen and the representation represents itself. And these two representations are one and the same experience.

A second reason why this is problematic stems from the first reason. If a representation cannot represent more than one thing at a time, then we would be continuously “reflecting” on our experience in order to know what it is that we are experiencing. We could not have an experience of the world while, at the same time, knowing that we are having an experience of the world. We would have to stop and reflect in order to know that we are having the experience of the world that we are having. Then we would have to stop reflecting in order to gain further experiences of the world, because, after all, we can only have one representation at a time and if we are currently focusing on the representation in our memory, then we could not be focusing on our representations of the world. There would have to be a continuous and cumbersome “jumping” back in for between our reflecting on our experiences of the world and our actual experiences of the world. But since this doesn’t happen, we can clearly have a representation that represents more than one thing at a time. So we can know that we are thinking while we are thinking.

There is also a problem with Costa's understanding of Descartes' position. Costa acknowledges that there may be knowledge of our thinking and knowledge of our existence in our periphery but not in our focus. Specifically, he states:

All of these things are in the fringes of the synchronic unity of my consciousness. It is correct to say that I'm *aware* of all these things, that I'm in a certain sense *grasping* all of them (if the air-conditioner stops, I will direct my attention to it). However, it is certainly not right to say that I'm *thinking* all or even some of these things, as Descartes would be forced to suggest. The only thing I'm consciously thinking are the words I'm writing, which are passing through the focus of my consciousness to lose themselves in the temporal extension of its diachronic unity. Consequently, Descartes give us a bogus example, which would be right only if he could show that we are in fact able to think simultaneously more than one thought in the sense of articulating these thoughts simultaneously in the focus or centre of our attention, when in fact we aren't able to do so (2001: 233).

In sections 2-2.2 I showed that he does not use the word "think" and "thinking" in the same way that we generally do. When we use the words we are usually referring to symbolic thought. But Descartes includes that act of perception and bodily experiences in his definition of being a "thinking thing". I even argue that we are best to think of it as 'experience' and rephrase Descartes thesis as 'I experience therefore I am' instead of simply 'I think therefore I am'. This means that knowledge and experiences that exist in our periphery or "fringe" of our conscious experience are a part of what Descartes would describe as thinking. Costa has misread Descartes.

5. Self Presenting Representations

So far I have been talking about the self referential part of representations; these are the part of the representations that refers to the “I”. Now I want to include self presenting representations. These are representations that represent themselves. For example, when I see the computer screen in front of me there is a representation of the computer screen and I know that there is a representation of the computer screen. Also, unlike self referential representations (which, I argue, are a part of every conscious experience), self presenting representations are not a part of *all* conscious experiences. It is possible to have a conscious representation of the world but not know that you have that specific representation of the world. However, despite the fact that self presenting representations are not universal, I believe that they are a part of most conscious representations. For organizational matters I will be dividing this chapter into two parts. The first part will discuss self presenting representations and our knowing that of experience, like the example of seeing a cup and knowing that you are seeing the cup. The second part will discuss self presenting representations in abilities and action. Such as reaching out to grab a cup and knowing that you are reaching out to grab a cup.

This section is important for two reasons. First, for Descartes, self presenting representations are a part of his theory of self representation and the ego. In order to give a full account of Descartes’ position on self representation, it is important to include self presenting representations in the discussion. Second, some confuse self presenting representations for self referential representations and, because of this, they argue that when there are no self presenting representations, then there are no self referential representations as well. But this is a mistake that needs clarification.

But before I get started there are a few other general distinctions between self presenting representations and self referential representations that I would like to introduce. A self referential representation is always a reference back to the “I” and it is the same kind of reference in every experience and action. Self presenting representations are different. You can have a self presenting representation of a visual experience, of a sound experience, an experience of thought, or an experience of action. Each self presenting representation represents a part of the “I”, but it does not represent the whole “I”. The self presenting representations are different for different experiences and actions. So there are multiple different kinds of self presenting representations but only one kind of self referring representations.

Also, as I have said above, it is possible for there to be a representation of the world without there being a self presenting representation. I think that these examples are rare, but they do exist. I will look at several examples where there seem to be conscious representations of the world without self presenting representations. However, I will continue to insist that even here there are self referential representations, because I have argued that there cannot be a conscious representation without a self referential representation (see chapter 3)

5.1 Self Presentation in Knowing that

Knowing that is distinguished from knowing how. Knowing how is when you are capable of performing an ability/action. Knowing that, on the other hand, is what you know about a state of affairs. For example, if I am seeing a colour, say, red, and I am conscious of my seeing the colour, then I know that I am seeing a colour. When it comes to consciousness, there is no such thing as perceptual experience in general; rather, every perceptual experience is a specific kind of experience. When I see, I always see shapes and objects and specific colours or black, white, and shades of grey. There’s no such thing as visual experience simple, with shapes and colours

added; rather visual experience is made up of these different kinds of visual experience. The same can be said for sound, taste, and smell. There is always a certain kind of sound, a certain kind of taste, and a certain kind of smell. There is no such thing as “perception simple” and then the kinds are added onto it; instead, the various kinds of perception are the experiences of perception.

Further, when a person has perceptual experiences, it does not take any extra thought to distinguish that an experience is taking place or to distinguish that a certain kind of experience is taking place. The event is immediately known by the person. If self representation was not an explicit part of the perceptual experience, then extra work would be needed for the person to identify the experience and the kind of experience.

The fact that every conscious experience is a kind of experience and we have immediate and direct access to this experience means that every experience of self presenting representation is a knowing-that. When I see a colour, like red, I know that it is me who is seeing the red or that it is me who is hearing a voice. And this self presenting representation must be explicit because if it was only implicit, then it would be possible for me to doubt who is having the visual experience. By this I do not mean that we would be able to doubt the accuracy of our visual experience. We can obviously do that. Rather, I mean that a person could doubt whether or not their own visual or auditory experience is actually their own. They could sincerely ask the question, is this visual experience that I am having my own? Or is it someone else’s? But this is unlikely. My own visual or auditory experience is obviously my own. And it follows from Descartes argument that this cannot be honestly doubted, indicating that the self presenting representation involved in the knowing-that of conscious experience are explicit.

But how widespread is self presenting representation? Most people will acknowledge that self presenting representation is a part of our conscious focus. But is it a part of other aspects of our conscious experience? Is it a part of our conscious periphery? To explore this I will take the example of driving a car. Sometime, when we drive a car, we get “lost” in thought. But we stay on the road and continue to obey traffic laws. Suddenly, our driving comes back into our conscious focus and we feel as if we do not remember anything from the last, let’s say, twenty minutes of driving. Are we conscious of our driving while we are lost in thought? And, if so, are we conscious in the self presenting representational way that is currently being discussed. This is a rather difficult question to answer, because it seems that we have no memory of the experience of driving and, therefore, we cannot report on the matter. This could lead some to argue that we are not conscious of our driving, or, if we are conscious, then we are not conscious in a self representational way. However, Ned Block, in conversation with Susan Blackmore (2006), discusses the phenomenon of potential “unconscious driving”. Block reports on “a pilot study in a driving simulator, in which they got people to space out and then probed them, asking ‘what are you experiencing now?’ And people always report the last ten seconds or so. So I think that in those cases there is a moving window of memory. I don’t think that’s a case where you’re having experiences but they’re not you” (2006: 31). If Block is right, then this is important to the topic at hand. First, it could indicate that you are conscious of the road and that this is not unconscious driving. Instead, what happens is that you are immediately forgetting your experience of the road. Probably because you are not focusing on it and it is not encoded into your memory. Second, it could indicate that you have the kind of self presenting representational consciousness that is currently being discussed, because you can report on the specific

experiences at hand. This could indicate that you have self presenting representations, even in your conscious periphery. Not just your center of focus.

However, even if there are self presenting representations in our periphery, it is possible that there are instances where you have a conscious experience of the world but you do not know that you have a conscious experience of the world. In other words, there may be conscious representation without self presenting representation. Possible examples of this are blindsight and inattentional blindness. In these situations the person does not know that they are having an experience of something. But when they are prompted to act on that thing, they can do so at a rate that is far better than chance. As I've already stated, I am willing to accept that self presenting representations are not a part of every conscious experience. However, I want to argue that self referential representations are a part of every conscious experience. So I need to be able to demonstrate that self referential representations are a part of these conscious experiences even though self presenting representations are not. The crux of my argument will be that people confuse self presenting representations with self referential representations. And think that just because self presenting representations are gone, self referential representations are also gone. This, I argue, is a mistake. I will use arguments given by Brook and Raymont (forthcoming) to help elaborate my point.

Brook and Raymont (forthcoming) acknowledge that self representation may be a part of most conscious experiences. But they argue that it is not a part of all conscious experiences. Specifically, Brook and Raymont argue that self representations are not *necessary* for a representation to be conscious. Brook and Raymont argue for this position from a theoretical standpoint and an empirical standpoint. I will address the theoretical argument first.

Brook and Raymont define a conscious representation with the use of the term 'appearing'. They build this definition on Nagel's (1974) definition of consciousness. Nagel states that a being is conscious if there is something that it is like to be that being. Brook and Raymont see this as a widely accepted definition of consciousness and they build their thesis in opposition to this definition. Nagel's definition includes self referential representation in it, because if there is something that it is like to be a being, then this focuses on the subject that is that being. However, Brook and Raymont argue that a revised version of this definition can be applied without self representational content. Specifically, they argue that the world can appear to a cognitive system without the cognitive system being conscious of itself. And if an object appears to a cognitive system then it is like something. This, Brook and Raymont argue, is conscious representation without self referential representation or self presenting representation.

But this leaves several questions to be asked. Is it possible for a person to forget that she exists while she is having this experience? Or can she mistake her own representation for someone else's? In regards to these questions I suspect that Brook and Raymont want to say, no. That she cannot forget her own existence and she cannot mistake her own representation for someone else's. But this means that, if self referential representation is not a part of the experience, then it is certainly close to hand and can be quickly retrieved before these questions become a real issue. However, Brook and Raymont can still argue that there isn't a self referential representation in *all* conscious experience while accepting that self referential representation is close to hand. Their main argument is that even if self referential representation is there, it may not be all that's needed for consciousness. So self representation is the base of consciousness, but it is not a requirement for consciousness.

But there is an objection to Brook and Raymond's position. All conscious experience of the world is an experience from a particular point of view and that the point of view can be seen as referring back to the subject who is having the experience. In chapter 3 I discuss how this point of view sets the "frame" for what we can see and what we cannot see. We are always aware of this "frame" of reference. We know that we can only see what is within it and we know that we can move our heads (or our whole body) to change it. It's impossible to have a visual experience of the world without also having a point of view. In a sense, it defines what is in our experience and what is outside of our experience. This would mean that all conscious representations have a self referential representation, because they explicitly include a point of view. To explain this idea, I will focus on vision and then expand it to other perceptual experiences.

Every visual experience is experienced from a position. We only see one side of an object and if we want to see the other side, we must either turn the object over or move our position. As we turn our head we change our point of view. When we stand up, sit down, or lie down we change our point of view. That a point of view is a part of your visual experience is the reason why you never wake up in the morning and wonder why the world is turned sideways. It's because the fact that you are lying down is a part of your point of view. Your peripheral experience also includes a point of view. The left side of your visual field is off to the left and the right side of your visual field is off to the right. To get this knowledge requires no extra experience. A child might not know the meaning of the words "left" or "right", but he or she is explicitly aware that he or she has to move in a certain direction in order to get the toy that is off to the left or right. These relational differences are an explicit part of the child's experience of the world. In fact, I don't know what a visual experience would be like without it having a point

of view. The idea that we could see all sides of an object, including the side that is sitting on the ground, is so strange that I cannot even conceive of its existence. So all visual experience has a point of view and this means that all visual experience has a self referential aspect to it. Of course, Brook and Raymond could respond by saying that, yes, there is a point of view and there is a self referential aspect, but it is unconscious. This is the position that is taken by Perry and Campbell in chapter 3. Since I've already addressed this position I will not say anything more. I will simply reiterate my argument. Self referential representation cannot be an unarticulated constituent because it plays a conscious functional role in action and thought. It also plays a role in the knowledge that we acquire. Unarticulated constituents can be misinterpreted. We can think that when a person says that "it's raining" they mean that it's raining in Ottawa, but they are actually talking about Calgary. This misinterpretation cannot take place with our visual experience and point of view. So self referential representations of visual experience cannot be unarticulated, they must be articulated. I will give further examples of this later on in this chapter.

The same can be said for the experiences of touch, sound, taste, and smell. When a part of your body touches something in the world (or even simply feels something, like the feeling of the wind), there is a point of view or self referential aspect to that feeling. It's rarely the whole body that has the experience but only a part of the body, like a hand, leg, or foot. You never touch or feel something but don't know what part of the body is experiencing the touch or feel. It's always a hand or a foot and so on. Further, that part of the body is always experienced in relation to the rest of the body. You feel something behind you or your hand is stretched out to the left of you. This means that all experiences of touch or feel are experienced with a point of view and have a self referential aspect to them. The same thing can be said for taste since the

taste is always in your mouth or mouth and nose. Sounds also have a point of view. You hear your friend call your name from the left or the right. And a smell is the smell of, let's say, those roses over there or of that dog. So these also have a self referential component. If there was no self referential component to smell then we would never be able to associate a particular smell with a particular object and another smell with another object.

However, it could be argued that while most experiences of the world have a point of view, not all experiences have a point of view. Sometimes we hear a sound or smell a smell and we don't know where the sound or smell is coming from. For example, if a person is deaf in one ear then the neurological processes that are responsible for giving our sound a point of view cannot work. In these cases our auditory experiences (or olfactory experiences) subordinate themselves to our visual experiences. A person who is deaf in one ear can still figure out who is talking to them and where the sound is coming from. They just need to associate the sound with a visual experience of a person moving their lips.¹ It's also important to note that you never believe that you are an infinite consciousness that is everywhere. So when you don't know where a sound or smell is coming from, you don't know this in relation to your point of view. In fact, the first thing you usually do is move your head and look around, thereby changing your point of view to see if you can pick out the source of the sound or smell. So even when you don't know where the perceptual stimulus is coming from, the experience still includes a point of view and, therefore, a self referential component.

So far, I believe that I have established that that all experiences of the world are experienced from a point of view. And that this means that there is a self referential component to the experience, even when there is no self presenting component. But I have only addressed

¹ This argument is borrowed from John Campbell (2002).

this issue from the conceptual side of the argument. Now I want to turn to the empirical side. Brook and Raymond focus on blindsight, inattentional blindness, and the dual route theory of vision to try and establish that there are conscious experiences without any self representational aspect. There are questions as to whether or not these cases are actually cases of conscious experience. But for the sake of the argument I will assume that they are genuine. I will address blindsight first.

5.1.1 Blindsight

The example of blindsight has been widely discussed in theories of consciousness. The phenomena can be described as follows. A subject has lost a section of their occipital or visual cortex, usually from a stroke or tumour, and this creates a blind-spot or scotoma in their visual experience. The loss is not experienced as a blank “spot”. It’s not like looking at a movie screen where a chunk of the picture missing. The area is simply gone from the experience. However, if the subject is put in a situation where a stimulus is presented in the scotoma and the subject is asked to guess, they will guess with an accuracy that is significantly above chance (Weiskrantz, 1988). There are also examples where the blindsight subject can act on the information when prompted by the experimenter. They do not always have a perfect accuracy and what can be guessed varies from subject to subject. To give an example, some (but not all) subjects discriminate colours (Stoering and Cowey, 1990). But the blindsight patients cannot engage on this information on their own. They require an external “indicator” for when to make their forced guess. In short, when forced to guess, the blindsight subject can guess significantly above chance, but he or she does not generally report any experience in the area and does not know if his or her guess will be correct.

Some argue that blindsighters are completely unconscious, but blindsight can also be interpreted in other ways. It could be argued that phenomenal conscious experience is required for us to be aware of information, then, if this is the case, blindsight could be a form of degraded content. And there is some evidence indicating that this is true. For example, if the subject is presented two visual stimuli, one in their normal visual area and one in the blind field, they will respond faster than when there is only one stimulus presented in their working visual field (Marzi, *et al* 1986). Also, Chokron *et al* (2008) writes about rehabilitation therapy with blindsight patients and has shown that improvement in laboratory experiments and self-reports in day to day endeavours is possible. The improvement varies from person to person and can include heightened sense of awareness that “something” is there, better accuracy in guessing, and a shrinking of the blind field. One participant (Patient #7) did not show any detectable improvement in the lab (although he did report some improvement in his daily activities). But this participant is the only one to have had the cortical area removed in surgery. This could indicate that visual recovery can take place only when there is an incomplete injury to the cortex. The authors explained these improvements through a theory called neuroplasticity. Neuroplasticity is the brain’s ability to change itself and create new interconnections between different regions. And despite the damage to the visual cortex many blindsight patients have neurological “islands” that are still in the damaged region. With practice, new and/or stronger connections can be formed between these “islands” and the rest of the brain. In patient #7 the surgical removal of the area did not leave the same “islands” as a stroke, which meant that fewer connections could be made. Most of the study focuses on functioning and informational access but the authors make one note about the phenomenological changes. They note that,

From a phenomenological point of view, at first, patients claimed that they did not see anything in their contralesional visual field and were ‘guessing’, then they reported that they ‘felt’ something in their blind field, they were then ‘pretty sure’ of their answer and finally they reported ‘seeing’ the stimulus in the ipsilesional visual field. Interestingly, this was also true for Patient #7 who experienced the same progression of ‘subjective’ improvement, despite the lack of improvement in the automated perimetry examination (Chokron, *et al* 2008: 314)

This indicates that the improved performance in accessing information is coupled with an increase of phenomenal experience. So this data also indicates that consciousness is playing a role in blindsight.

To be sure, there were still significant limitations in the blind field. The subjects do not report to have an experience of visual events but a “sense” or “impression” that something is there. The authors speculate that this “may be a function of reflexive orienting and may influence their ability to localize and discriminate at above chance levels” (317). And the improvement was best under certain types of situations, especially ‘emergency’ conditions, which may also indicate an improvement in awareness of mild and usually unnoticed reflexes. For example, Patient #7 reports that he was able to catch a falling razor even though it was in his affected area. The authors described this increased capacity as the ability to exploit “implicit” residual detection to aid in conscious detection.

However, the patients did not recover their phenomenal visual experience of what is in the blind field. There are reports of “‘seeing’ the stimulus” but not experiencing the object. They had better awareness of their surroundings and reported feeling more comfortable moving about

in their daily activities. Although most of these improvements are still limited to the degree of informational access found in blindsight, it is coupled with an increased phenomenal awareness of mild reflexes. This increase in ability could be viewed as an increase in conscious awareness that is similar to a degraded peripheral conscious experience. But the person has enough information that they can include it in a degraded but similar manner as our peripheral awareness. It starts to “go along” with their conscious goals, but it does not have any significant detail. And the experience of the visual reflexes is also similar to peripheral experience. There is a reflexive consciousness that allows the person to access some information, but without an awareness of a particular thing or object. The degraded phenomenal experience is such that they cannot focus on it and bring it into the center of their conscious attention, because this requires significantly more phenomenal awareness than the blindsight patient can muster. But the blindsight patients (above) can be described as having improved their degraded peripheral visual experience.

This seems to be a clear example of a representation of the world without a self presenting representation of the world. Even the subjects who have gone through rehabilitative treatment do not have knowledge that they “see” the object. They just have a sense that something is there. So perhaps the rehabilitative subjects have regained some self presenting representation, but are still missing most of it. But this does not mean that there isn’t any self referential representation in the blindsight subject. It is still possible that the blindsight subject has self referential representation without self presenting representation. If this is a genuine experience of the object, then it is an experience from a point of view. The subject sees the object from a certain position. The subject can only see one side of the object, and the subject can act on the object from his or her certain position. The fact that the subject can reach out his or her

hand, orient it in the right manner and grasp the object, indicates that the subject has a conscious experience of the object from his or her position and that a reference to his or her position is a part of the experience of the subject. If there was no self referential representation in the experience of the object then the subject would have difficulty in acting on the object. He or she would fumble the object and possibly miss it altogether.

So I argue that this fact, that the subject is capable of acting on objects even though he or she cannot focus on the object, implies that there is an explicit self referential representation in this degraded peripheral experience. The fact that the subject can act on the object shows that the representation includes a point of view and, as I said above, this point of view includes a self referential representation. Also, we should note that the subject is conscious of the fact that he or she does not have a self presenting representation of the object. The fact that the subject is conscious of this implies that he or she is conscious of herself or himself and what he or she is (and is not) capable of consciously grasping from the visual experience. Some researchers have even noted that the patients who perform the best are also the ones who are judged to be the most “passive” (Weiskrantz 1988). But if there was no self referential representation in the degraded peripheral experience then the opposite would be true, because the patients would have to try and focus on their blind region in order to acquire this information and be able to act on the information. If the person has no self referential representation in the conscious experience (even this degraded conscious experience), then the person would not act accordingly, because the self would be completely unaware of the information. Self referential representation is required for action. Without it the person does not know how to act in relation to an object. So the fact that people do best when they let their periphery work on its own can be seen as evidence for self referential representation, even in this degraded peripheral experience.

5.1.2 Inattentional Blindness

Inattentional blindness is when a stimulus is presented in a subject's visual field, but the subject does not claim to see the image. I will first give a small historical introduction (including the most famous experiments), then I will move onto the examples that Brook and Raymond focus on in their argument for conscious representation without any self representation.

Inattentional blindness was first reported by Neisser (1979) when he asked participants to watch two different groups playing a ball-passing game. One group is wearing black and the other group is wearing white. The participants are supposed to track the ball from one of the groups and ignore the other. During the experiment a woman strolls through the scene carrying an open umbrella. But a significant number of subjects do not see the woman; they are too busy trying to keep track of their group's ball. Simons and Charbris (1999) perform a similar experiment, but with a person in a gorilla suit walking through the scene. Once again, a significant number of subjects completely miss the gorilla. Other examples of inattentional blindness are when a subject is asked to concentrate on a fixation point and is asked to note a feature of an introduced object on or near that point. After a few trials a second object is introduced in a slightly different spot. But this second object is masked by 1.5 seconds of masking. At least one quarter of the subjects report that they do not see the second object.

But is the object of inattentional blindness consciously experienced? Brook and Raymond point out that there is some evidence to suggest that it is. For example, Mack and Rock (1998) show evidence of significant priming in examples of inattentional blindness. When a subject is asked to complete a word stem (a word stem is when only part of a word is provided and the subject is asked to fill in the rest of the word) they are significantly more likely to complete it to match the word they have been primed for, even though they sincerely claim not to have seen the word. Further, Moore and Egeth (1997) have shown that the stimulus can affect perceptual

processes—such as grouping of elements in the visual field into objects—and it can affect a person’s actions on the environment, even though the person claims to not have seen the object.

There is another problem with this theory, namely that it could be a rapid forgetting and not a problem with perception. Specifically, if the conscious experience is a peripheral experience then the experience is there, the person has it, is aware that they have it, but, because they do not focus on it the experience isn’t encoded into the person’s memory and they forget about it. If the reader will remember from section 4.1, when Block discussed the possibility of “unconscious” driving, and argued that it is possibly a rapid forgetting of the peripheral experience and not actually “unconscious” driving, it is possible that the same can be said for inattention blindness. And the priming and effects on action can be accounted for by saying that there isn’t a *complete* forgetting, just a partial one. There is some evidence for this: Most et al (2001) show that features that are more similar to the task that the subject is asked to perform are more likely not to suffer from inattention blindness. Lavie (1995) has argued that the inattention blindness is due to the perceptual overload that takes place when we are engaged in an action. The less engaged we are, the less likely there is to be inattention blindness. Of course, some might find it implausible that a person would forget that a gorilla walks across a screen while people are playing a game a catch. But, on closer inspection, this isn’t too difficult to imagine. The person sees the gorilla but is so focused on what they are doing that they never turn their attention to the gorilla and it isn’t encoded into memory. A person could deny that they would forget something as obvious as a gorilla, but when you compare it to the so-called “unconscious” driving phenomena, it doesn’t look that farfetched. After all, what’s more likely, that we see the road while we drive down the highway but do not encode it into our memory and forget about how we stayed on the road? Or that we drive unconsciously for twenty minutes

while (miraculously) we manage to stay on the road? So it's just as possible that the subjects of inattention blindness consciously see the object, have self-referential representation of the object, and even (possibly) have a self-presenting representation of the object, but they are so focused on the task of counting the passes of the balls, the object of focus, or distracted by the masking, that they do not encode the object into their memory and forget about it. An everyday example of this forgetting is all the sights and sounds that enter our conscious periphery throughout the day, but, because we do not focus on them, we forget them within seconds of them leaving our periphery.

There is another problem with the interpretation of inattention blindness as representation without self-referential representation. When directed, the subject is capable of some action. This means that a point of view is a part of the experience. If the subject can reach out and direct her arm and hand in the appropriate way, then there must be a self-referential part to her experience. Once again, it seems that there is a mistake between the ability to symbolically think about the stimulus and vocally report the stimulus and self-referential representational content. Self-referential representational content doesn't exist only in linguistic thought. There is self-referential content in vision, hearing, tasting, smelling, and touching. In fact, all conscious experience has a point of view and, therefore, all conscious experience has self-representational content. If the person can act with their body in response to a stimulus, then the experience of the stimulus includes a point of view. And if the experience includes a point of view then there is a self-referential aspect.

Brook and Raymond try to demonstrate that inattention blindness doesn't have any self-representation (be it self-referential representation or self-presenting representation). However, a closer analysis demonstrates that there is certainly self-referential representation and possibly self-

presenting representation as well. I think that it is best to view inattentional blindness as a forgetting than a seeing without knowing that we are seeing. If this is the case, then there isn't just a self referential representation, there could also be a self presenting representation. It's just that the experience isn't focused on and is forgotten. Similar to how we can drive a car while daydreaming. There is an experience of the road, but, because we are focused on something else it isn't encoded into our memory and is forgotten. Brook and Raymond's third example has more to do with action than knowing that, so I will address it in the next section.

5.2 Self Presentation in Action

Before I start my analysis of self representation in action, I would like to talk a bit about the overlap between knowing how and knowing that. All acts of knowing how contain knowing that, each involves the other. For example, if I know that I am seeing two different colours then I know how to distinguish between the two colours. Also, if I know how to ride a bike, then I know that I push with the pedals, sit on the seat, and steer with the handle bars. Instead of sitting on the handle bars and trying to steer with the pedals. But knowing how (or ability) cannot be reduced to knowing that's. Just because I know all the knowing that's that make up the action of riding a bike, this does not mean that I know how to ride a bike. I would probably fall to one side or the other the first time I got on the bike. So knowing how takes practice and cannot be eliminated from the philosophical lexicon.

In section 5.1 I discuss self representation in knowing that. Now I want to focus on self representation in knowing how or action. In most actions there seems to be the two kinds of self representations being discussed: self referential representations and self presenting representations. As I type on the computer I am both conscious of the keyboard as my fingers hit

the keyboard (the representation of the keyboard) and I am conscious of the fact that my fingers are hitting the keyboard (the self presenting representation of my fingers hitting the keyboard). Also, my fingers are a part of me, they have a relation to the rest of my body, and they represent a part of me. So there isn't just a self presenting representation, there is also a self referential representation.

However, it is important to note that I do not seem to be conscious of all the action that takes place while I am typing. For example, my fingers move quickly, I don't really know how fast my fingers are moving and I don't really know how they hit each key in the order that they do. It simply comes from practice that I can type at the speed that I am typing. Indeed, the best way for me to mess up and have to slow down my typing is for me to try and concentrate on the subtle movements of my fingers. Our fingers seem to go along with the top down goal of our desired actions, but we do not consciously control the specific movements of our fingers.

Another example is when I reach for my coffee cup. I am conscious of the fact that I want my coffee cup and I am conscious of the fact that I am reaching in a certain direction, but I do not have a self presentational consciousness of the exact angle of my reaching and the exact movement of my fingers as they prepare to grasp the cup. Jeannerod (1986) gives a detailed study of the movement of the fingers in grasping. He found that the fingers adjust during the reaching component of the action, long before the fingers touch the object. However, what's important is not just that the fingers move into grasping position before they reach the object but they do so in such a way that that action includes a representation of the object.

Some might want to argue that there is not conscious representation at all, but I think that this is a little too hasty. It is possible that we have a conscious representation of our fingers movements while we type or reach for an object, but we do not have a self presenting

representation of our fingers movements. The reasons for this would be the same reason for accepting a representation (without self presenting representation) in blindsight. If blindsight can be a representation without a self presenting representation, then there is no reason to say that there is no conscious representation here either. Another reason for thinking that this might be the case is that we seem to have some control over our action. If we want to stop or deliberately close our fingers so that the grasp will be too small, we can do so. This is in contrast to actions that we have no control over. Saccadic eye movements could be a good example of a completely unconscious act, because we have no control over them. They happen, we do not know when they happen, we cannot stop them from happening, and we cannot make them happen. This seems to distinguish saccadic eye movements from the fine tuning movements of the fingers as I type on the computer or reach for my coffee cup. The saccadic eye movements are completely unconscious whereas the fine tuning movements of the fingers are conscious, but there is not self presenting representation of them. So we are conscious of the fine tuning movements but we are not conscious of the fact that we are conscious of them. Also, there is evidence for a self referential representation in these types of actions. The fact that the fingers adjust appropriately to grasp the object shows that there is both a representation of the object and a representation of the agent that is acting in relation to the object. Paulignan et al (1996) studied the orientation of the axis defined by the position of the index finger and thumb during that act of grasping the same object at different placements. Even in simple tasks where arm motion was that only factor, they found that the finger axis remained invariant to the body axis of the subject, which indicates that the action was affected by a body centric (or ego centric) frame of reference. So even the fine tuning of the fingers in action have a self referential representations in the action.

5.2.1 The Dual Route Theory of Vision

The theory that there are two visual routes for perception and action is the last example Brook and Raymond give to try and demonstrate that a person can see the world without any self representation (be in self referential representations or self presenting representations). The theory states that vision has two routes for processing visual information. One is a route for perception, which allows a person to know what they are seeing. The other is a route for action, which allows the person to act on what they are seeing. The route for perception is in the 'ventral stream' and the route for action is in the 'dorsal stream'. Of course, the 'action route' requires some knowing that and the 'what route' requires some knowing how, so these routes are not evidence for a disassociation of the philosophical concepts of knowing that and knowing how. We should not confuse the psychological distinction for the philosophical distinction. An example of the psychological distinction is elaborated by Goodale et al (1991), a 35-year-old woman who suffered irreversible damage to the lateral occipital region and the parasagittal occipitoparietal region. She demonstrated poor perception of shape or orientation, which, through testing, could not be explained by simple sensory deficit. 15 months after her accident there was still dissociation between the patient's ability to perceive objects' orientation and her ability to direct accurate reaching movements toward objects in different orientations. To test her orientation perception, they asked the patient to choose which card angle was orientated in the same way as a slot and to indicate the orientation of a block placed on a table. The patients' performance on these tasks was significantly impaired. She would even make such significant errors as judging horizontal to be vertical. In contrast, the patient still had the ability to act on the information. When the patient was asked to reach out and put the hand-held card into the slot her performance was excellent. Video recording showed that the patient began to properly orient the card immediately after starting to move her hand, demonstrating her anticipation that the card

needed to be a certain angle and that it is not just a trial and error action. Also, when asked to pick up a block on the table the patient oriented her hand appropriately very early in the action, indicating that she knew how to pick up the block. A variety of tests were performed leading Goodale et al to conclude that the patient's "difficulty seems to be in using visual orientation information for perceptual or cognitive purposes, even though such information is accurately used in visuomotor action" (1991: 155). Other patients have demonstrated the opposite ability. They can use information for perceptual and cognitive purposes, but they cannot act on the information.

Goodale and Milner (1992) give a summary of some of the evidence for the dual route thesis for perception and action. Originally, the theory argued that the distinction between 'what' and 'where'. But Goodale and Milner (along with others) argue that the distinction is better described as the distinction between 'what' and 'how'. They point out that there are patients who are unable to recognize faces, pictures, or abstract designs, even though they can make their way through the world with ease. And there are patients who can recognize objects and designs, and the orientation of designs, but cannot act on the information. A closer look indicates that these subjects do not just have difficulty discovering the position of objects (the 'where' thesis) and reaching in the right direction. They also have difficulty positioning their fingers or adjusting their hand when reaching toward an object that can be oriented at different angles. This would indicate that it isn't just a question of 'where' but also a question of 'how'.

Goodale and Milner also argue that this distinction can be found in normal subjects. They refer to studies of eye-hand coordination during visually guided aiming. For example, a subject will not report noticing a change in object location if that change takes place during saccadic eye movement, even in forced choice guessing. However, manual aiming in action shows near

perfect adjustment in hand/arm movement during action. Examples like these seem to indicate that there is a distinction in normal subjects as well as disabled subjects.

When it comes to consciousness and self representation, Goodale and Milner (1992) interpret the data in a way that is different from Brook and Rayment. Specifically, Goodale and Milner argue that when the subject can see the object and can cognitively think about the object then the subject has a conscious experience of the object. They also note that there is a 'viewer-centered' aspect to the experience, because the subject sees the object from a particular point of view. However, when the subject cannot see 'what' the object is, but can act on the object Goodale and Milner argue that the information is largely, if not completely, unconscious. The reasons for this is because most of the fine adjustments that go on in action, such as the positioning of fingers and the fine tuning of hand movements, are thought to be unconscious actions.

However, there is a problem with this interpretation. First, the subject has a self representation of the action itself. The subject knows that he or she is partaking in the action and knows which direction and how he or she is acting. The 'fine tuning' is unconscious, but the entire action is not unconscious. Second, just because the subject cannot say 'what' is being seen (in regard to position and angles), this does not mean the subject does not see anything at all. There is a difference between the representation of an object and a self presenting representation of an object. Given that some of the subjects have the deficit of 'what,' there is a clear indication that the subject does not have a self presenting representation of the object. But this is different from their being no conscious representation at all.

This leads Brook and Raymont to argue that, in the cases where there is an ability but not a knowing that, the subject has a representation of the object but no self representation at all. If by self representation they mean self presenting representation then I agree. But if they mean self referential representation then I disagree. The representation of the object clearly has a point of view, because the subject is able to act on the object. The subject knows which direction to move his or her hand and knows how to position his or her fingers. This indicates that the subject has a reference to the self in his or her representation of the object. In the last section I gave evidence for self referential representations in the fine tuning of finger movements in actions. If they have self referential representations, then it would seem that the actions addressed in this section also have self referential representations. Of course, Brook and Raymont could argue that the self referential aspect of the representation is there but unconscious, but I have already addressed why this cannot be so in chapter 3 when I discussed the theories of Perry, Anscombe, and Campbell. Basically, the ability of the subject to act on the object demonstrates that the subject has a representation of the object that includes information about where the object is in relation to the subject, this means that there is still a self referential representation. Even though there is no self presenting representation.

One of the main problems with Brook and Raymont's position is that it deals with controversial examples. As the names imply, most people think that blindsight, inattentional blindness, and one of the dual routes in vision are examples of unconscious information processing. It could easily be argued by someone that this evidence isn't evidence for conscious representations *at all*. And there are people who make this argument. Brook and Raymont disagree. But they have tried to argue that there is only evidence for a representation of the object. What I point out is that *if* the evidence that Brook and Raymont give is actually evidence

for a conscious representation of the object, then it is equally evidence for a conscious self referential representation. The evidence supports conscious self referential representation just as much as it does conscious representation of the object. There is no reason to reject one but not the other. So *if* one is conscious, then there is equal reason for supposing that both are conscious.

It is a mistake to think that when there is only ability, there isn't any self referential representation. I believe that I have clearly demonstrated that all actions, even actions that do not include a self presenting representation, have a self referential representation. In the next section I will examine how all three representations work together in our interaction with the world.

5.3 Self representation

I have addressed three kinds of representations: First, there are the uncontroversial representations of the external objects. Second, there are the self referential representations of the self that are a part of all conscious experiences (knowing that's) and actions (knowing how's). Third, there are the self presenting representations that are a part of most experiences of the world and most of our interactions with the world (however, they are not a part of the fine details of our actions). Now I would like to demonstrate how these three kinds of representations work together in the workings of consciousness. I will do so with examples. I will start with the simplest example first, a teenage boy who previously couldn't see colours, but, with the aid of special contact lenses, was able to see colours. This example is one where we are addressing self representation only in the knowing that and the knowing how of immediate conscious experience. Then I will move onto a more complex example of Sally the wine taster, where we have knowing that's of experience and knowing how's of action. Then I will finish with the most complex example, how a batsman in cricket hits a ball. This example has knowing that's of experience, knowing how's of action, and unconscious processes.

I want to briefly look at an example of a person who was colour blind and then regained his colour vision. Kevin Staight was a teenager who has a rare eye defect that caused him to see the world as black and white until he was fitted with a special contact that enabled him to see colours. This gave Kevin new experiences and new knowing that's. But it also gave him new abilities, the ability to distinguish between colours that he couldn't see before. In a newspaper article he said:

After I put them on I went for a walk and slowly saw the world in colour for the first time. Up until then, I didn't have any idea what colour was because I couldn't see it. I couldn't stop crying because the world looked so different from what I was used to. The reds just kept on jumping out at me and I had to ask my grandparents which colours were which because I didn't have a clue.

It has opened up a whole new world for me. I never realized just how beautiful things like trees and flowers are (Cited in Tye 2009: 125)

The things to take note of in this quotation are as follows: First, Kevin knows of his own existence and has a self referential representation that he is the one having these new experiences. Second, as Kevin goes for his walk, colours such as red "jump out" at him and, therefore, Kevin knows-that he is seeing various different colours and he knows that it is he who is seeing various different colours. Third, Kevin is able to point out the different colours and ask his grandmother their names so Kevin knows-how to distinguish between the various colours. So Kevin has a self presenting representation in both his experience of the different colours and his ability to distinguish between the different colours.

That was a fairly simple example. Now let's try a more complex example. Let's take the example of someone, let's say Sally, who is learning to be a wine taster. First, Sally has probably read up on the tastes of various different wines. But just because she has knowledge by description about the different kinds of wines does not mean that she is able to distinguish between two kinds of wines in a taste test. She is still a beginner so the two wines taste the same.

First and most obviously, Sally knows that she exists. Her knowledge of her own existence is not simply discovered in the tasting of the wine but is consciously experienced in order for the wine tasting to begin. This self referential representation is not just in her act of lifting the wine glass to her lips, although it is in that too, but it is also in the very knowledge of having a taste experience. And when she has the taste experience this further reinforces her knowledge of her own existence. In short, without this self referential representation, Sally would (and could) never choose to be a wine taster.

Second, when Sally tastes the wine she knows-that it has a certain taste. This certain taste is immediately obvious to her. She knows it the moment she tastes it. Further, she knows that it is her who is having this certain taste and this is also as obvious to her as the taste itself. So she also knows this immediately, there is no extra step required in order to figure out that it is her who is experiencing the certain taste.

Third, Sally knows-how to distinguish this taste from other tastes. Given that the taste is immediately obvious to her, if she has the experience of two different tastes the fact that she has two different tastes is also immediately obvious to her. So she knows-how to distinguish between the two different tastes. However, since Sally is a beginner wine taster she also knows-that the

two wines taste the same to her and that she currently doesn't know-how to distinguish between the two wines.

But as Sally keeps practicing her perceptions in taste will enhance. She will begin to taste the subtle differences in the wines and when this happens, she will know-that one wine tastes slightly like raspberries and she will know-that the other wine tastes slightly like cherries. The very moment Sally acquires the knowing-that's of these two wines these taste experiences will constitute a knowing-how to distinguish between the two wines. Sally will acquire the knowing-that and the knowing-how immediately and she will also immediately know that it is she who as acquired this new knowledge. Of course, Sally might say something like, "I can't believe I just did it!" But this is a figure of speech and not a literal statement. Sally knows that she did it.

To show how these three representation work together in action, I will give a detailed description of Sally's distinguishing the taste of two wines. Sally sees the cup. She has a conscious representation of the cup. Sally knows that she sees the cup. She has a self presenting representation of her seeing the cup. Sally also knows about her own existence and that it is her who is seeing the cup. So she has a self referential representation of her seeing the cup. Then Sally reaches for the cup. Sally is conscious of the fact that she is reaching for the cup. She still has a representation of the cup. She also has a self presenting representation of her reaching for the cup. She has a conscious representation of the details of how her fingers move as she reaches for the cup. And, since these details are body centered, she has a self referential representation of the details of how her fingers are moving as she reaches for the cup. But she does *not* have a self presenting representation of the details of how her arm and fingers move as she reaches for the cup. Then Sally grabs the cup and lifts it to her mouth. She has both a visual representation of the cup and the representation of the cup by her hand. She is conscious of her bringing the cup to her

mouth and touching the cup against her mouth. She also has a self presenting representation of her touching the cup against her mouth, even though she does not have a self presenting representation of all the fine details that goes on as she touches the cup to her mouth. But Sally does have a self referential representation of both the general act of touching the cup to her mouth and the fine details in action that are required to touch the cup to her mouth. If she didn't have this self referential representation, then Sally would probably miss her mouth. As she drinks, Sally has a representation of the wine (by taste and feel) in her mouth. She has a self presenting representation of the wine (by taste and feel) in her mouth. She both knows that she has wine in her mouth and knows that the taste is a certain kind of taste. Specifically, she knows that it is a kind of taste that is different to the last taste of wine that she had. And she has a self referential representation of the wine in her mouth (by feel and taste). So the representation of the objects, the self presenting representations of knowing that's and abilities, and the self referential representation that is a part of all conscious experience interact with each other to guide her knowledge and interaction with the world.

The last and most complex example is studied by Land and McLeod (2000). They analyze how a cricket batsman hits a ball. The problem of how a batsman hits a ball is an interesting problem for self representation. The bowler throws the ball at high speed with a variety of different possible angles; the ball bounces at a variety of different possible positions, and then leaves the ground at a variety of new angles and speeds. Of course, a batsman does not hit the ball on every try, but a good batsman will hit the ball on some occasions. Land and McLeod compared the methods used by three batsmen of different skill levels. They set the ball speed at fast to medium-fast speed for cricket (25 m/per/s and up) and analyze the eye movements of three batsman with a head mounted eye camera. Land and McLeod report that

even the best batsman takes about 200 ms to adjust his or her swing based on visual information. So the batsman has to decide his or her swing angle approximately two thirds into the ball's trajectory.

What's interesting is that we have all heard the statement, "keep your eyes on the ball", but the best batsmen do not follow the ball all the way in. Instead, the eye saccades onto the ball as it leaves the bowler (or in this case, the "pitching machine") and then, after following the ball, their eyes saccade to the approximate place where the ball will hit the ground. Then they eyes move to slightly below the trajectory of the ball after the bounce, approximately where the ball will be when the batsman can make his or her last adjustments in their swing. So there is a technique in hitting the ball that includes automatic (unconscious) saccades, representations of the ball, self presenting representations of the ball, and self referential representations. Land and McLeod also report that this eye saccading (where the player does not keep his or her eye on the ball) is also found in table tennis and baseball. This pattern was seen in all three batmen, regardless of their skill level. However, there were differences between the three batsmen. And these differences reflect their abilities. The best batsman showed more pursuit tracking than saccade "jumping" in eye movements (pursuit tracking is when you follow the ball instead of saccade). So he had a longer representation of the ball, self presenting representation of that representation, and self referential representation of himself in relation to the ball. The second best showed the second most pursuit tracking and the amateur batsman showed the most saccade tracking. The amateur was also the slowest in response time for "finding" the ball as it came out of the pitching machine. The camera showed that this is a result of slower saccade initiation. So the amateur has less time to track the ball from the beginning. The good batsman and the expert batsman showed no difference in this first saccade, indicating that their differences in

performance rest in the expert performing more pursuit tracking than saccades. This difference in pursuit tracking allowed the expert more information and, therefore, a greater ability to self-adjust his swing. Most of the balls in the test were hit, indicating that it is not necessary to keep one's eye on the ball all the way through. Instead, it is best to fix on the delivery, jump to the point where the ball will bounce, and then to where the ball will be up to about 200 ms before it reaches the bat.

Land and McLeod focus mostly on the eye movements of the batsman, but I want to add a hypothetical analysis of how self-representation can play a role in hitting the ball. This example is the most complex example, because it has both conscious and unconscious aspects to it. But if we walk through the process step by step, then we can see how self-representation plays a role. First, there is a saccade that fixes the eye on the ball coming out of the pitching machine. All three participants had this saccade, but the amateur was the slowest. The saccade is unconscious, the participants have no control over how quickly or slowly it happens. Instead, it demonstrates how practice can make unconscious reflexes faster, which makes the more practiced players better at spotting (and therefore hitting) the ball. Once the eye is fixed on the ball there is clearly a representation of the ball. But there is also a self-presenting representation of the ball and a self-referential representation. The batsman sees that the ball is coming at him at a certain angle and that the ball is moving in a certain relation to him. That the more practiced player saccades faster to this point of view gives him or her more time to acquire information about the ball's speed and trajectory and this allows him or her more time to adjust his or her swing. Then there is a saccade to the point where the ball will bounce. This saccade is also automatic and unconsciously performed. However, it is performed based on the conscious information received from the first conscious representation of the ball. At this point there is a representation of the ball bouncing, a

self presenting representation of the ball bouncing, and a self referential representation of the ball bouncing in relation to the position of the batsman. The batsman doesn't just see the ball bounce, he knows that he is seeing the ball bounce. And he knows that the ball is bouncing at a certain trajectory in relation to him. This is why the information is useful for hitting the ball. If there was no relation to the batsman, no self referential representation, then the batsman would not be able to use this information to judge how he or she should swing the bat. After the ball bounces there is a third unconscious saccade that jumps to where the ball will be approximately 200 ms after the bounce. It is from this conscious representation, self presenting representation, and self referential representation, that the batsman makes his or her final adjustments on how to move his or her body and swing the bat.

I'd like to briefly touch on the differences between the participants' levels of expertise. Since the amateur is the slowest to focus on the ball, he has less time consciously representing the ball. He also has less time being aware of the representation of the ball and knowing where the ball is moving in relation to him (the self referential representation). This means that he has less time to gather information about the ball's speed and trajectory and he has less time to adjust his actions in relation to the ball. The difference between the intermediate player and the expert player shows a similar pattern. They both saccade at the beginning at approximately the same speed, but the intermediate was more likely to saccade between positions and the expert displayed more pursuit tracking of the ball instead of saccades. This means that the expert also had a longer conscious representation of the ball. And during this pursuit tracking he also has a self presenting representation and self referential representation of the ball coming towards him. This means that the expert also has more information about the ball's speed and trajectory in relation to himself. This gives the expert more time to adjust his body and swing, making it more

likely that he will hit the ball. So even though a batsman does not keep his eye on the ball the whole time, the longer he keeps his eye on the ball the more information he receives and the better he is at hitting the ball. And self representations play an important role in the batsman's adjustments of the body and swing.

Now that I have discussed representations, self referring representations, and self presenting representations I would like to apply this analysis to a classical philosophical problem, how we gain knowledge of the external world and how we confirm/disconfirm our knowledge of the external world.

6. Our Knowledge of the external world

The purpose of this chapter is not to solve problems, but to address some problems that can be taken up at later dates. Specifically, I hope to demonstrate that self referential representation can play a role in the explanation for how we gain knowledge of the external world (even though I will not be solving the problem itself). Also, I hope to briefly touch on the question of perception by looking at one theory (Campbell's) of disjunctive perception. Disjunctive perception tries to solve the problem of the external world by arguing that the external world is the constitutes of our conscious experience. If this is so, then Cartesian representations do not exist. Then I will briefly address mild scepticism to our knowledge of the external world and outline how such a theory may work. The point of this chapter is only to give preliminary remarks of some problems that could be worked on in the future.

How we gain knowledge of the external world is a classical problem in the philosophy of mind. One that Descartes was keenly aware of. Of course, Descartes' solution to the problem was to postulate the existence of God, declare that God is all good and, therefore, cannot be a deceiver. Then Descartes' argues that we accurately perceive the external world because God allows us to accurately perceive the external world. However, this "solution" will not satisfy most philosophers today. Instead, we need a solution for how we gain knowledge of the external world and how we can confirm or disconfirm that knowledge. In other words, we need a method for determining whether or not our experience of the world is actually an experience of a real object that exists out there beyond our experiences, or whether our experience is simply a hallucination.

Before I lay out my own theory of how we gain knowledge of the external world, I will introduce two different attempts. First, I will introduce Ismael's theory from situated cognition. This will help me further elaborate on the problem itself, while also helping me elaborate on why self referential representation needs to be involved in the solution to our knowledge of the external world. Then I will address John Campbell's Relational View of perception to give an example of an attempt to solve the problem. Campbell critiques Cartesian representationalism for the Relational View, arguing that the problem of how we have knowledge of the external world is not actually a problem. By showing that Campbell's critique of Cartesian representationalism is wrong and the Relational View does not solve the problem, I will establish that the problem of how we gain knowledge of the external world is a real problem. The goal of this chapter is not to solve the problem of how we gain knowledge of the external world. It is simply to show that self referential representation needs to be involved in the solution.

To be clear, I am a mild external world skeptic. I believe that we can distinguish between dreams and reality, because reality has a continuity that dreams do not have. I also believe that we can realize that mild illusions are illusions. However, we cannot determine whether or not we are being fooled by an evil demon or whether or not we are a brain in a vat. These types of hallucinations (where hallucinations validate previous hallucinations) are impossible for us to uncover.

6.1 Ismael and Situated Representation

An example to help elaborate the problem of our knowledge of the external world is as follows: We have two worlds, A and B. A and B are epistemologically separate. World A *seems* to get a message from world B, but how does A know that the message came from B? And how can A verify that the message came from B? Further, even if A chose to guess that the message

came from B, A needs to respond to B to let B know that A got the message. A tries to respond to B. Now, how does B know that the message came from A? Remember A and B are epistemologically separate. It looks like the message came from A in B's world, but B has no access beyond B's world. B could guess and try to respond to A and tell A that B got the message. But how does B know that A got the response? After all, any *apparent* response from A is just as ambiguous as the initial message. This can go on for some time, so I will stop here. But we can think of the problem of our knowledge of the external world in these terms. We have two "worlds". A is the world of our internal mental representations. And B is the external world. There seems to be no way to confirm that a message coming from B is actually coming from B. And if A tries to respond there seems to be no way of knowing that the response to B actually made it to B.

To address this issue, Ismael adopts the perspective of situated cognition. Loosely put, situated cognition states that the mind (and self) is situated in the world. But this is not just a claim that there is always something beyond our perceptual experiences, behind us for instance. Properly argued, situated cognition is a paradigm shift on how we view the mind. The mind is no longer separate from the world, as it is in the traditional Cartesian view. Instead, the mind is embedded in the world and a part of the world. So, to refer to the example above, world A is embedded in and a part of world B. So there is really only one world. Ismael argues that this is what makes possible our knowledge of the external world.

But there really does *seem* like there are two worlds. After all, there are the internal representations of the world and there is the external world itself. Ismael (2007, 28) compares this to the two worlds of Celsius and Fahrenheit. Here it also seems like there are two "worlds" when it comes to temperature. I know Celsius because I grew up in a country that has adopted

the metric system. I know nothing about Fahrenheit, it's foreign to me. But there is really only one world here. And the same can be said for a situated view of the mind. Also, Ismael points out that what we need is representational overlap to move from one kind of knowledge to another. If I know that Celsius and Fahrenheit overlap at -40° , then I can gain some knowledge of Fahrenheit from Celsius. When Celsius is -40° then Fahrenheit is also -40° . But in order for me to truly gain knowledge of Fahrenheit I need two representations. Knowing that Celsius and Fahrenheit overlap at -40° is a good start, but I can't know what Fahrenheit is when Celsius is 10° without a second representation. As it turns out, Fahrenheit is 50° when Celsius is 10° . From this I can conclude that whenever Celsius goes up 50° from -40° , Fahrenheit goes up 90° . From this I can conclude that for every degree that Celsius goes up, Fahrenheit goes up $9/5^{\text{th}}$ of a degree. Now I have complete knowledge of Fahrenheit from my knowledge of Celsius. Ismael argues that situated representation works the same way.

However, there is a problem with Ismael's argument. It doesn't give us knowledge of the external world. If we have two representational fixed points, both of these fixed points would be in the same world, it would be equivalent to having two points in Celsius but no points in Fahrenheit. What she wants to argue is that if we have a "fixed point" that aligns with a point in the external world, then we can use situated cognition to gain knowledge of the external world. But she never explains how we get that external fixed point in the first place. She presupposes that we have the fixed point referent in the external world, without explaining how we get it. But if we already had a fix point referent in the external world then we would already have a way of gaining knowledge of the external world (it would be how we got that fixed point referent). So this argument cannot be used to give us knowledge of the external world, because it presupposes a solution without giving a solution.

Despite this problem, Ismael's view has some advantages. Once we find a way to gain knowledge of the external world, we still need the two representations in order to "translate" our knowledge. So, once we have solved the problem, we need two representations to properly claim that we can have knowledge of the external world. However, Ismael adopts the theories argued by John Perry and Anscombe, claiming that there is no internal self referential representation. Instead, there's (1) the representation of the object and (2) the self presenting representation of that representation. But if we follow the arguments on blindsight, inattentional blindness, and the dual route theory of vision in section 5-5.3, then we have to accept that a person can have at least some knowledge of the external world without a self presenting representation. A person suffering from blindsight does not have a self presenting representation, but he or she is capable of guessing the correct answer at greater than chance averages. Sometimes he or she is even capable of acting on the object in the "blind" area. A person in a state of inattentional blindness also has a representation of the object without a self presenting representation. But this representation affects semantic processing and affects the person's actions. And the dual route theory of vision also implies that a person can have knowledge of the external world, and act on that knowledge, without a self presenting representation. As I believe I've shown in my arguments against Perry and Anscombe, they do not give a satisfactory reason for rejecting (internal) self referential content in conscious experience. And I believe that my analysis of blindsight, inattentional blindness, and the dual route hypothesis demonstrates that there are self referential representations in these conscious experiences. So it makes more sense to argue that Ismael's first representation is the representation of the object and her second representation is our self referential representation and not a self presenting representation. If two representations are needed for our knowledge of the external world and we can have knowledge of the external

world without self presenting representations, then self referential representations are the only representations left to take the role as the second representation.

To help further this theory, it will be useful to further elaborate on Ismael's position. Ismael states that we do not have a self referential representation of the self. Instead, we "find" the self by situating the self in the world. In a sense, we work from the outside in. We perceive the world and find ourselves in the world from this perspective. She compares it to finding oneself on a map in the mall. You walk up to the mall in search for the red dot marked "YOU ARE HERE". First you see one store, then another, you start to recognize some of the stores that you see, and finally you find the dot and exclaim, "Here I am!" It's an interesting thesis, but I disagree. You never really forget where you are. You always know that you are "here", wherever here happens to be. Of course, there are times when you're walking or driving, perhaps through a new neighborhood or city, and you stop and think to yourself, "where am I?" But in these situations you haven't lost yourself. Instead, you've lost your position in the world. To compare it to Ismael's map example, when you are lost you still (and *always*) have the red dot saying, "YOU ARE HERE" in your hand (or, to be more accurate, in your head). What you need to do is find the right spot to *place the dot on the map*. Once you've found the right spot, then you know where you are *in the world*. But, once again, you never actually lost yourself in this process.

The fact that you never lose yourself, but you do lose your position in the world, can also be interpreted as evidence for a self referential representation in conscious experience. If Ismael were right, then we really could lose ourselves. We could mistakenly take someone else's current conscious experiences for our own. We would have to learn about our own point of view in comparison to other people's point of view. And we would have to take some care in making sure that it is actually me that is thinking. Indeed, it would be appropriate to say that "it thinks"

instead of “I think”. But, as I have argued throughout this paper, this is a mistaken position. We do not lose ourselves and have to find ourselves, even if we do lose our position in the world. We do not mistakenly take other peoples’ current conscious experiences for our own. We do not have to learn that our own visual point of view is our own and not other peoples. And we do not need to take care to make sure that it is actually I who is thinking. That is why it is appropriate to say that “I think” instead of “It thinks”.

6.2 Campbell’s Critique of Cartesian Representationalism

Campbell (2002) is interested in how we acquire knowledge via demonstratives. For example, if I say that ‘all dogs are furry’ (which has no demonstratives in it), then we may not need knowledge of the external world. But if I say that ‘that dog is furry’ (which has the demonstrative ‘that’ in it), then we need to have knowledge of the external world. Specifically, we need to have knowledge of which dog is being talked about. Campbell argues that consciousness plays a role in our understanding via demonstratives, because we need to be conscious of the object that the demonstrative refers to in order to have knowledge about the demonstrative’s reference. I have no problem with the main argument of Campbell’s thesis. However, Campbell also argues that we need to reject Cartesian representationalism in order to gain this knowledge of demonstratives reference. Instead, Campbell argues that we need to adopt a relational view of consciousness, where the objects themselves constitute the contents of conscious experience. This is where Campbell and I part ways. I will briefly outline Campbell’s critique of the Cartesian representational thesis, showing that while his solution seems to work at first glance, it doesn’t align with what we find in practice. I will leave my own solution for the next section.

As I said, Campbell is interested in how we gain knowledge via demonstratives. But what exactly does he mean by this? Suppose that you walk into a room that you have never been in before. You see a couch, a chair, a table, a TV, and pictures on the wall. Now, do you see the couch, the chair, the table, the TV, and the pictures on the wall or do you experience a representation of them. Campbell (2002: Chapter 6) argues that if we experience only a representation of them, then there is no way for us to distinguish between our experience of hallucinations and the external world, because the hallucination is indistinguishable from the representation of a real object.

In the representational view, there is no experiential difference between hallucinations and real objects because the conscious experience consists on being in a certain internal representational state and might have nothing to do with our external surroundings. Of course, representationalists will argue that it is your surroundings that usually cause the representational state, but will acknowledge that this representational state can be caused by something else, such as a drug or chemical imbalance in the brain. So the experience could be the same.

In contrast, the relational view argues that there is an experiential difference between your hallucinations and real objects. This is so, because, according to Campbell, the contents of experience are not constituted by an internal representational state, rather, the contents are constituted by the external objects. And when there are no external objects to constitute the experience, the experience will be different. This view is sometimes called the 'disjunctive' theories of perception; because it argues that there is a disjunction between hallucination and the real thing. The representational view does not recognize this disjunction, it says that one's experiential states can be the same regardless of whether or not it is a hallucination or the real thing. This is what causes the problem of our knowledge of the external world. So Campbell

argues that the representational view needs to be discarded and the relational view should be adopted to explain how we acquire knowledge of demonstratives and the external world more generally.

However, there are problems with Campbell's conclusion. It seems to work at a first glance, but it does not work in practice. It's fine to say that the experiences would be different, but, in reality, the experiences of illusions and hallucinations are no different from the real thing. Campbell doesn't give any examples of how the experiences are different. Instead, he discusses the brain in a vat, claiming that if you were a brain in a vat, your experiences would be different from real experiences of the external world. But when we look at examples from real life, things turn out differently. Take, for example, the Müller-Lyer Illusion below. We all know this illusion. The line segment on the left looks longer than the line segment on the right. But they are actually the same length. However, this figure looks just like it would look if there was a longer line segment on the left than the right. We know that this is an illusion. We don't even want to believe that the line segment on the left is longer than the right. But we still see the line on the left as being longer than the right, so there doesn't appear to be a disjunction between the experience of mild illusions and an experience of two lines of different length. If there were two lines of different length, it would look that same as the illusion below.

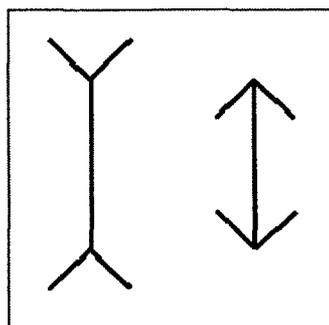


Figure 1: The Müller-Lyer Illusion

What can be said for mild illusions can also be said for extreme hallucinations. The hallucination is indistinguishable from the real thing. Being a person who has had an episode of psychosis, I know from firsthand experience that a major hallucination is indistinguishable from the real thing. When you have a hallucination of a voice, it 'sounds' just like a real voice. It's so real that you are firmly convinced that it is a real voice. The experience is so real that you even have trouble letting it go after you have been put on medication. This problem of letting it go is so common that therapists will address the issue of re-interpreting your memory so that you recognize that your past experiences were actually a hallucination. If they don't do this it is very easy to continue on in life believing that you have (or had) special insight into the world and have been spoken to by God/the devil/ a spirit. If the relational view of consciousness were correct then people suffering from psychosis or schizophrenia should be able to tell that they are having hallucinations and are not experiencing the real thing. But this is clearly not what happens. Most (if not all) people suffering from psychosis or schizophrenia believe the delusions and hallucinations that they are experiencing. For them they are indistinguishable from the real thing. So the relational view has to deal with examples that do not align with its claims for disjunctivism.

Another example of major hallucinations is dreams. When we dream, we have all sorts of conscious experiences. We see things, hear voices, and interact with the "world". Further, we generally experience rather strange phenomena. The world doesn't work according to the ways that we are used to. People can fly, you come into contact with your dead relatives, and you mysteriously move from one situation to the next with no congruency. But if our hallucinatory experiences are fundamentally different from our accurate representational experiences, as Campbell's relational view claims, then we should be able to know that we are dreaming. But

when we dream we don't know that we are dreaming. We go along with the dream world as if it were the real thing. It is only when we wake up that we realize that we were dreaming. And think to ourselves, "what a strange dream I just had."

There is another problem with the relational view of consciousness, which has to do with the claim that the constituents of conscious experiences are the external objects we see, hear, tastes, touch, and smell. If this were true then everyone should have the same experience. But this is not the case. We all agree that two people can taste the same food, but only one person may like it. And the same can be said of smell. Even in vision there are differences in our experience of the external world. For example, one person can look at an object and say that it is pure red. But another person can look at the same objects, in the same light, and say that it is clearly a reddish orange. The colour vision of people varies, and it varies according to gender, race, and age. People can look at an object and see the object as having different colours (Block, 1999). To give a specific example, a study of 50 males showed that they divided into two broad groups, with the difference between the groups traceable to a polymorphism in the L-cone photopigment gene (Winderickx et al, 1992). But if the objects are the constitute of our conscious experience, then one group is hallucinating, because they are not seeing the "proper" colour, but which group? If we are going to argue that colour is a property of the external object and not a part of an internal representation, then we should be able to measure this external property to figure out who is hallucinating and who is correct. If we have a disagreement on what the height of an object is, then we can take out a ruler and measure the object to see who is right and who is wrong, and if both are wrong, we can see who is closest. All properties of external objects have an external (objective) method of measurement, but colours, tastes, sounds, feels, and smells do not have this external method of measurement. The relational view is in the

awkward situation of having to argue that these properties of conscious experience are actually the properties of external objects, but they do not work like the rest of the properties of external objects. I consider this view to be indefensible. If Campbell is going to claim that external objects are the constituents of our experience, then he needs to be able to explain what it is about the external objects that constitute the various different aspects of our experiences. The same can be said for all other aspects of consciousness that are experienced during hallucinations or dreams. They cannot be explained away and they cannot be reduced to an external property. So they must be an internal representation.

So the relational view falls short. But even if we adopt the representational view of consciousness, we still seem to be left with the problem about how we acquire knowledge of the external world. How is it that we know that we are not dreaming? How do we know that we are not suffering from a grand delusion like a person with schizophrenia? How do I know that the computer in front of me is not a hallucination? In the future I would like to look at other theories of disjunctive perception, while also elaborating on a more explicit theory of perception of my own. But these preliminary remarks help layout some problems that need to be addressed. Now I will give an argument for how we can confirm or disconfirm our experiences, thereby allowing us to have knowledge of the external world, with mild external world skepticism.

6.3 On Hallucinations

If Campbell's argument for how we gain knowledge of the external world is inadequate, then how do we decipher between hallucinations and real experiences. First I want to make a general statement about my argument. I am not giving a purely logical argument for how we can decipher hallucinations from real world experiences. Instead, I am giving a practical argument. An argument that is not a perfect method of distinguishing between hallucinations and real world

experiences, but a method that, nonetheless, is capable of confirming or disconfirming the representational experience. I believe that this a common sense response to the question of how we know that we are not hallucinating. I take it that everyone—or at least almost everyone, there is probably a few adamant skeptics in the world—will agree that they are not hallucinating most of the time. And when they wake up, they can tell the difference between a dream and the real world that they wake up to. So everyone has a way of telling hallucinations from reality. If they didn't, then they would have never figured out that the Müller-Lyer illusion is actually an illusion. How we confirm or disconfirm our experience of the external world is by *acting* on the representational information. If the action is a success, then we know that we are not hallucinating. If the action is a failure, then we have strong reason to believe that we are hallucinating. Further, even if the action is thought to be a success, because of a further hallucination, there will be consequences that cannot be hallucinated. Let's take a simple example. You see a dagger in front of you. You wonder why it's there, floating about, suspended in space. You reach out to try and grab the dagger. But your hand goes right through it. The fact that you've attempted to act on the information, but failed in that action, is evidence that you are actually hallucinating. However, if we have a hallucination of a suspended dagger and try to act on that hallucination by grabbing it, but, instead of disconfirming our hallucination, we confirm our hallucination and think that we have 'grabbed' the dagger, then further disconfirmation will demonstrate that it is a hallucination. If we try to stab someone with the dagger, the person will not be injured. If we try to cut an apple with the dagger, the apple will not be cut. If we try to cut ourselves with the dagger, we will not be in real pain. So, even if there is some hallucinatory conformation, there will be a later disconfirmation against the hallucination, because there are disconfirmations that cannot be hallucinated. If, on the other hand, an evil demon makes it so

that we have nothing but confirmations to our hallucinations (the apple cuts, for example), then there is no way to disconfirm our hallucinatory experience. And there is no way to tell that it is a hallucination.

Now let's turn to the examples from the last section. Take the Müller-Lyer example. A first year student, who has never seen the Müller-Lyer illusion, will look at the illusion and believe that one line segment is longer than the other line segment. But when you ask them to act on it and confirm this belief, say, by measuring the lines with a ruler, they find out that the two line segments are actually the same length. By acting on the illusion and trying to confirm the illusion, they learn that it is actually an illusion. They adjust their beliefs accordingly and act differently the next time they see the illusion.

But what about our knowledge of our dreams? That seems to be a little harder to explain. After all, most people do not realize that they are dreaming *while* they are dreaming. Usually, we know that we were dreaming because we wake up and find ourselves in our bed (or perhaps our couch). This dramatic disjunction in experience is not usual for waking life. It is the empirical evidence that causes us to re-evaluate our previous experiences and rightly interpret them as a dream. Small children are more susceptible to believing in the reality of their dreams. Even when they wake up, especially if it is an emotionally charged dream, like a nightmare, they need reassurances that their past experiences were really only a dream and nothing more. But over time, with successive experiences and confirmations, most people learn to recognize dreams as only dreams and learn to disassociate them from reality. Of course, there is the controversial phenomenon of lucid dreams. I don't want to digress too far into this debate, but from what I understand, in a lucid dream a person comes to realize that they are dreaming and is capable of controlling their dream. This too relies on acting on the dream experiences, realizing that one can

do things that one cannot usually do, and this evidence confirms (or possibly disconfirms) the persons belief that he or she is actually dreaming. The reason why we can distinguish dreams from reality is that dreams are incoherent. One minute we are talking to a dead relative in a restaurant, the next minute we are in the forest, the minute after that we are flying. The fact that there is not a continuous coherent experience in our dreams, but there is a continuous coherent experience in your life while we are awake, allows us to distinguish between dreams and being awake.

Now for the much more difficult question of psychosis and schizophrenia, why is it that these people never realize that they are having hallucinations? This example is much more complicated, but I think the answer has to do with the fact that there isn't just one hallucination, instead, there are multiple. And these multiple hallucinations (including the hallucinations of rigidly held beliefs) really do seem to confirm the previous ones as the person interacts with the world. These delusions—be it a delusion of reference from the television set (a delusion of reference is when you believe that you are getting special messages or your are interpreting the plotline or words coming from the television as having special significance to you), voices that comment on your thoughts, perceptions, and actions, or paranoid delusions of being persecuted by forces beyond your control—all seem to confirm each other. They coalesce into one giant “grand narrative” that the person finds him or herself in. And as the person acts out these delusions, more delusions seem to continue to confirm previous delusions. So the method that we usually use to disconfirm hallucinations, like the Müller-Lyer illusion and dreams, would work against the person suffering from psychosis or schizophrenia. What they need is a disconfirmation of the delusions and hallucinations, but it rarely and possibly never comes. However, there are some hallucinations that even a person with psychosis can uncover. If the

person continuously hallucinates that he or she is eating but continues to lose weight, then the losing weight disconfirms the hallucination of eating. If a person believes that the apocalypse is coming tomorrow and the apocalypse doesn't come, then this disconfirms the hallucination. But those people who always think that we can tell between hallucinations and reality, they need to explain why people suffering from psychosis or schizophrenia do not realize that they are hallucinating. The fact that these people do not realize this, confirms my position of mild external world skepticism.

Of course, the person suffering from psychosis or schizophrenia is hopefully put on medication and, if the medication is successful, then this brings their delusions and hallucinations under control. They start to lose their rigidly held beliefs about persecution and they start to get disconfirmations against their hallucinations. Slowly but surely, they reinterpret their past experiences as being irrational, not the way the world actually works, and improbable. Further, the very act of being on medication and seeing that the medication works, along with the diagnosis of a mental illness, pushes the person along in their reinterpretation of past experiences. In fact, some people with schizophrenia still have some symptoms, like continuing to hear voices, even when they are on medication. But the medication controls the hallucinations enough to greatly enhance their ability to confirm or disconfirm the hallucinations with reason and actions. This allows them to put the hallucinations aside and live relatively normal lives.

The basic premise for how we decipher a hallucination from the real thing is by acquiring further evidence, via further experiences, of confirmation or disconfirmation, especially through action. I acknowledge that this method is not perfect. For example, people who are in the grips of a psychotic episode cannot use it. But I think that it accurately reflects how we go about confirming and disconfirming hallucinations, such as dreams, in the real world. We can tell when

we wake up from a dreaming state, and we can disconfirm hallucinations in most cases. Some people might not like this theory, because it does not provide an irrefutable *proof* for how we have knowledge of the external world. But to argue that we can have irrefutable knowledge of the external world contradicts the evidence. People are sometimes taken in by illusions, like the Müller-Lyer illusion, and people suffering from psychosis and schizophrenia are more or less helpless to their hallucinations. So anyone who wants to argue that we have an irrefutable method of gaining knowledge of the external world and distinguishing between hallucinations/delusions and real experiences, needs to explain why these examples are out there. The fact that my theory give a method for distinguishing between illusion, hallucinations, delusions, and the external world, but acknowledges that this method is not perfect and can sometimes lead us astray is an advantage and not a shortcoming. It accurately depicts the reality of our situation.

7. Conclusion

The main purpose of this essay was to explore the universality of self referential representations. Then, in chapter 6, I give an introductory analysis of the role of self referential representation in solving the problem of how we gain knowledge of the external world. We started with a look at the historical origins of the theory with Descartes. In our quick analysis of Descartes we found that he does not actually argue for a Cartesian theatre of consciousness, but, instead, he argues that the perceptual and bodily experiences are also a part of the I that thinks. This means that Descartes does not separate representations of the external world from the ego. Instead, these representations are a part of the ego. This means that the self referential part of conscious representations that Descartes expresses in the proposition 'I think therefore I am' are a part of all conscious experiences, not just symbolic thought. Descartes also argues that self referential representations are a part of our abilities and actions, since we also experience them. This is the part of Descartes view that I accepted, tried to expand on, and defend.

After rethinking Descartes I turned to the arguments of Perry, Anscombe, and Campbell. These three have argued that self referential representation is either unconscious or non-existent in our experiences of the world. Perry argues that it is an unarticulated constituent; Anscombe argues that there is no reference for the word "I", and Campbell takes these two thesis and argues that the word "I" is a logical function like "and". I started with Perry, showing that self representation does not work like unarticulated constituents, because there is no possibility for misinterpretation, error, and no need for learning. Then I turned to Anscombe and argue that since Kaplan has distinguished between pure indexicals, which have fixed references, and pure demonstratives, which have a reference that is not fixed, Anscombe's argument that the word "I" does not have a reference is no longer valid. "I" is a pure indexical. Specifically, it always refers

to the speaker or thinker. This reference is fixed in this way in all situations. And since it is impossible for there to be speaking or thinking without a speaker or thinker, this means that “I” always has a reference. Campbell tries to treat the “I” like a logical function, but the “I” doesn’t work like a logical function, because it is context sensitive. You cannot create universal rules for its use or draw truth tables. This means that the “I” cannot be a logical function and must be a part of the proposition that has a reference in order to determine the truth value of the proposition.

Next I tried to demonstrate that the most basic aspect of self referential representations, our knowledge that we exist, is incorrigible. I look at three arguments from Armstrong and one argument from Costa. Armstrong’s arguments work well against the incorrigibility of other mental knowledge, such as whether or not one is in pain, but they do not work well against my more rolled back version of incorrigibility. Costa’s argument runs a foul on two fronts. First, he misinterprets Descartes by thinking that Descartes’ definition of “think” only includes symbolic thought. Second, he mistakenly believes that a representation can only represent one thing at a time, once these mistakes are sorted out Costa’s argument falls apart.

Then I introduce self presenting representations and address Brook and Raymont’s arguments against a self referential representation being a part of all conscious experience. Brook and Raymont argue that the world can appear to the subject in such a way that the subject is only consciously aware of the world and not herself. However, I point out that the subject has a point of view in its awareness of the world, and that the point of view is actively involved in framing and constraining the subject’s experience of the world, so this awareness of the world necessarily includes a reference to the position of the subject and, therefore, the subject herself. Brook and Raymont also argue that blindsight, inattentional blindness, and the dual route theory of vision

give evidence that a person can have an experience of the world without any self representations. If by self representations they mean self presenting representations, then I agree. But if they mean self referential representation, then I disagree. I analyze some empirical evidence on blindsight finding evidence that the experience that the blindsighter has includes a point of view and the self awareness that the subject does not have a self presenting representation of the representation and, therefore, there is a self referential aspect to the experience. Also, the fact that the blindsighter can sometimes act on the experience further demonstrates that there is a point of view. In inattentional blindness I argue that what is actually happening is a forgetting and not a representation without self representation. And in the dual route theory of vision I argue that this must also include a point of view, because the person is capable of acting on the information. I also analyze the evidence in action, demonstrating that our actions, even the fine tuning in our hand/arm gestures, includes an ego-centric reference, which I interpret to be evidence for a self referential representation in the action. At the end of the chapter I give three real world examples of how self representation plays a role in our behaviour: there is the example of Kevin Staight, the teenager who could not see colours until he was fitted with special contacts; the example of Sally learning how to be a wine taster; and the example of how a batsman hits a ball.

In the last chapter I give an introductory analysis of self representation in the problem of how we gain knowledge of the external world. I look at Ismael's argument from situated cognition and find that it presupposes us having knowledge of the external world and doesn't explain how we got that knowledge in the first place. However, despite this setback, I argue that Ismael's theory that two references are needed in order for us to "translate" between the external world and the representational world is a valid position. So I argue that we cannot have only a

representation of the object, but we must also have a second representation, the self referential representation, to explain how we can acquire knowledge of the external world. On delusions, I argue that we confirm or disconfirm our experiences by trying to act on the experiences and gather further empirical data on the experiences. We know that a dream was a dream, because there is a serious break in the flow of experience when we wake up and find ourselves in our bed. We know that a hallucination of a dagger is a hallucination when we try to grab it and our hand goes right through it. But a person suffering from psychosis or schizophrenia does not know that they are having delusions because their further delusions seem to confirm their past delusions. My argument for how we gain knowledge of the world is not a perfect logical argument. Instead, it is an argument based on our real world actions, confirmations, and disconfirmations of experiences. According to my argument we can be led astray, but I consider this to be a positive instead of a negative, because in the real world we are sometimes led astray.

There are several directions that further research could take. First, in this essay I only spent a small amount of time on Descartes. The reason for this is that I only wanted to use him as an introduction to the theory of self referential representation. However, a more detailed discussion of Descartes could be useful. Specifically, why he does not actually adopt a Cartesian theatre perspective and why self referential representations are a part of all conscious experience instead of just symbolic thought would help further clarify Descartes position from misinterpretation.

On the topic of whether or not self referential representations have no referent (Anscombe), have an unconscious referent (Perry), or are an interesting mixture of the two (Campbell) could use a further elaboration. Both Perry and Anscombe are influenced by statements made by Wittgenstein. But they are influenced by different sections of his work. I

chose to put Wittgenstein aside because such an exegesis would take me too far off the main discussion in this work. But a longer work (or a separate essay) could go back to Wittgenstein, work through his different positions, and see which positions are still valid today.

Second, on the topic of incorrigibility, I stepped around the issue of self representing zombies. Some (most notably Chalmers) may have objections to this. A more thorough examination of the zombie argument would be useful. Specifically, an investigation into how we know that we are not zombies seems to be an interesting question. The zombie argument claims that consciousness is epiphenomenal. But if we truly know that we are conscious (we know that we are not zombies), then it seems that consciousness must have at least one causal power. It causes us to know that we are conscious. But then the epiphenomenal thesis is false. After all, if consciousness doesn't cause us to know that we are conscious, then it is possible for us to have this "knowledge" even though we are not actually conscious. This would be an example of a self representing zombie. But this would mean that we don't actually know whether or not we are zombies. This seems to be a rather peculiar conclusion; some might even argue that it is a *reductio ad absurdum* against the zombie/epiphenomenal thesis of consciousness. But this argument needs to be spelled out in much greater detail and can help further the thesis that self referential representation is a part of conscious experience.

Also on the topic of incorrigibility, there is a very long history of argument for and against incorrigibility. Some argue that nothing is incorrigible, while others argue that all first person conscious states are incorrigible. I chose to take a middle route, arguing that self referential states are incorrigible and, because of this, I only addressed arguments that were made directly against self referential incorrigibility. However, the other theories of incorrigibility and the critiques against them could deserve some attention in a longer essay.

Further, while I am satisfied with my discussion of blindsight and Inattention blindness, a further elaboration could have been made on the dual route theory of vision. There is an especially large amount of data on this theory, with lots of interesting issues to examine. I only discussed enough to try and demonstrate that self referential representation is a part of these conscious experiences. But I would like to perform a much broader discussion on the dual route theory, especially on its implications for the philosophical distinction between knowing that and knowing how. But this would have been a digression from my analysis of self representation. So I put it aside for a later date.

In summary, there is much more that can be said on the topic of self referential representation, self presenting representations, and the role they play in consciousness. But I believe that I have established a firm grounding for the thesis, one that makes clear distinctions, backs them up with arguments, and analyzes their roles in a variety of circumstances.

Bibliography

- Anscombe, G.E.M. (1975) *The First Person Mind and Language* Guttenplan, Samuel (ed.) Oxford University Press: London UK
- Armstrong, David M. (1963) "Is introspective knowledge incorrigible?", *Philosophical Review*, 72: 417–432
- Block, N. 1999. Sexism, racism, ageism and the nature of consciousness, in R. Moran, J. Whiting, and A. Sidelle, eds., *Philosophical Topics* 26 (1&2), *The Philosophy of Sydney Shoemaker*. Fayetteville: University of Arkansas Press
- Brook, Andrew and Raymont, Paul (forthcoming) *A Unified Theory of Consciousness* Chapter 1. The concept of consciousness, the elements of a theory.
- Blackmore, Susan (2006) Ned Block. *Conversations on Consciousness* Oxford University Press: New York
- Campbell, J. (2002) *Reference and Consciousness*. Oxford University Press: New York
- Campbell, John (2004) What Is It to Know What 'I' Refers to? *The Monist* 87 (2) pp 206-218
- Clark, Andy (1997) *Being there: Putting brain, body, and world together again*. MIT Press: Cambridge, MA
- Costa, Claudio (2001) 'I'm Thinking' *Ratio* 3: 222-233
- Chokron, S. Perex, C. Obadia, M. Gaudry, I. Laloum, L. Gout, O. (2008) From blindsight to sight: Cognitive rehabilitation of visual field defects. *Neurology and Neuroscience* 26 305-320
- Dennett, Daniel (1991) *Consciousness Explained* Back Bay Books/Little Brown and Company: New York
- Descartes (1641/1988) *Selected Philosophical Writings* trans. Cottingham, John; Stoothoff, Robert; and Murdoch Dugard. Cambridge University Press: Cambridge UK
- Goodale, M. A., Milner, A.D., Jakobson, L.S., and Carey, D. P. (1991) A neurological dissociation between perceiving objects and grasping them. *Nature* 349: 154-156
- Goodale, M.A., Milner, A.D. (1992) Separate visual pathways for perception and action *TINS* 15: 20-25
- Ismael, J.T. (2007) *The Situated Self*. Oxford University Press: New York
- Jeannerod, M. (1986) *The Neural and Behavioural Organisation of Goal-Directed Movements*.

Oxford University Press: Oxford, UK

- Kaplan, David (1989) *Demonstratives. Themes From Kaplan*. Almog, Joseph, Perry, John and Wettstein, Howard (eds.) Oxford University Press: Oxford UK
- Land, M.F., McLeod, P. (2000) From eye movements to action: how batsmen hit the ball. *Nature Neuroscience* 3: 1340-1345
- Lavie, N. (1995) Perceptual load as a necessary condition for selective attention, *Journal of Experimental Psychology: Human Perception and Performance*, 21, 451-468
- Lewis, D. (1988) What Experience Teaches. *Proceedings of the Russellian Society*, 13: 29-57
- Mack, A., and Rock, I. (1998). *Inattentional Blindness* MIT Press: Cambridge, MA
- Marzi, C. A., Tassarini, C., Aglioti, S., and Lutzemberger, L. (1986) Spatial summation across the vertical meridian in hemianopics: A test of blindsight. *Neuropsychologia*, 35, 867-880
- Most, S.B., Simons, D.J., Scholl, B.J., Jimenez, R., Clifford, E., and Chabris, C.F. (2001) How not to be seen: The contribution of similarity and selective ignoring to sustained inattentional blindness *Psychological Science*, 12, 9-17
- Nagel, T. (1974) What is it like to be a bat? *Philosophical Review*, 83: 435–450
- Neisser, U. (1979) The control of information pickup in selective looking. *A tribute to Eleanor Gibson* A.D. Pick (ed.) Hillsdale, NJ
- Perry, J. (1979) The Problem of the Essential Indexical. *Noûs* 13 (1) pp. 3-21
- Perry, J. (1986) Thought without Representation. *Proceedings of the Aristotelian Society, Supplementary Volumes* 60 pp. 137-151
- Perry, J. (1990) Self-notions *Logos: Philosophical Issues in Christian Perspective* 1 pp 17-31
- Pauligana, Y., Frak, V.G., Toni, I., and Jeannerod, M. (1997) Influence of object position and size on human prehension movements, *Experimental Brain Research*, 114: 22634
- Rees, G., Russell, C., Firth, C., and Driver, J. (1999) Inattentional blindness versus Inattentional amnesia. *Science*, 286, 849-860
- Rolands, M. (2009) Situated Representations. *The Cambridge Handbook of Situated Cognition*, Robbins, Philip and Aydede, Murat (eds.) Cambridge University Press: New York
- Simons, D.J., and Chabris, C.F. (1999) Gorillas in our midst: Sustained Inattentional blindness for dynamic events. *Perception*. 28, 1059-1074

Stoerig, P., and Cowey, A. (1990) Wavelength sensitivity in blindsight. *Nature*, 242, 916-918

Tye, M. (2009) *Consciousness Revisited* MIT Press: Cambridge Massachusetts

Weiskrantz, L. (1988) Some Contributions of Neuropsychology of Vision and Memory to the Problem of Consciousness. *Consciousness in Contemporary Science*. Oxford University Press: New York.

Winderickx, J., Lindsey, D. T., Sanocki, E., Teller, D. Y., Motulsky, A. G. & Deeb, S. S. (1992) Polymorphism in red photopigment underlies variation in colour matching. *Nature* 356(6368):431-33.