

Dennett's Verificationism: Attack on Cartesian Materialism

by

Brandon Michaud

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

© 2016

Brandon Michaud

Abstract:

In this thesis I explore the role that verificationism plays in Dennett's attack on the implicit (and sometimes explicit) view of consciousness that he titles Cartesian materialism. I begin by tracing the lineage of verificationism from its historical roots in the Enlightenment period up to and at its heyday in logical positivism. I then proceed to argue that Dennett's position on content involving the intentional stance produces a form of realism that is too weak. While I disagree with Dennett's approach to intentional content's status as a manifestation of the intentional strategy, I find myself largely in agreement with his eliminative materialist and verificationist approach to qualia.

Acknowledgements:

I would first and foremost like to thank Professor Andrew Brook for all of the helpful, patient and stimulating guidance he has given me while writing this thesis. Professor Brook's work has been a source of great inspiration to me and has in many ways inspired my approach to philosophy and the topic I ultimately chose to work on.

I would also like to thank the committee who oversaw and commented on my thesis. The questions and comments were very helpful in developing my considered positions.

I dedicate this thesis to my parents, my grandmothers, my sisters, as well as my grandfathers (the latter of whom are unfortunately no longer with us).

Table of Contents:

<i>Introduction</i>	1-3
<i>Chapter 1 Introduction</i>	3
<i>The Enlightenment</i>	3-7
<i>Hume & Kant's Proto-Verificationism</i>	7-11
<i>Logical Positivism/Empiricism</i>	11
<i>The Vienna Circle</i>	11-12
<i>Elimination of Metaphysics</i>	12-13
<i>The Verification Criterion</i>	13-14
<i>Verification in Practice and in Principle</i>	15-17
<i>Strong and Weak Verificationism</i>	17
<i>Internal Criticisms</i>	17-19
<i>An Alternative to Verificationism</i>	19
<i>Am I Arguing for Verificationism?</i>	20
<i>Conclusion</i>	20
<i>Chapter 2: Intentionality and</i>	20-21
<i>Instrumentalism in Folk Psychology</i>	
<i>(Introduction)</i>	
<i>What is Folk Psychology?</i>	21-22
<i>The Three Stances:</i>	22-26
<i>Physical Stance</i>	22-23
<i>Design Stance</i>	23-24
<i>Intentional Stance</i>	24-26
<i>Dennett's Instrumentalist</i>	26-28
<i>View on Folk</i>	

<i>Psychology</i>	
<i>Real Patterns, The Intentional Stance, And Interpretationalism</i>	28-29
<i>To What Extent is Dennett Plausibly A Realist or Not?</i>	29-36
<i>Eliminative Materialism & Industrial Strength Realism</i>	37-39
<i>Folk Psychology as a Theory</i>	39-41
<i>Is Folk Psychology a Theory?</i>	41-45
<i>Where Do Dennett and I Agree And Disagree?</i>	45-46
<i>Conclusion</i>	46
<i>Chapter 3: Two Models of Consciousness:</i>	46-47
<i>Cartesian Theatre and Dennett's Multiple Drafts Model/Fame in the Brain</i>	
<i>Cartesian Materialism: A Model of Consciousness</i>	47-48
<i>Discarding Cartesian Dualism</i>	48-49
<i>The Cartesian Theatre & Cartesian Materialism</i>	49-50
<i>Multiple Drafts Model of Consciousness</i>	50-55
<i>Conclusion:</i>	55
<i>Dennett's Attack on Qualia</i>	55-57
<i>Are Qualia Ineffable?</i>	57-62
<i>Are Qualia Intrinsic?</i>	62-65

<i>Are Qualia Private? Are Qualia</i>	64-68
<i>Incorrigible?</i>	
<i>Phi Phenomenon and</i>	68-71
<i>Verificationism</i>	
<i>Metacontrast</i>	71-72
<i>Analysis</i>	72-74
<i>Conclusion</i>	74-75
<i>Conclusion of Thesis:</i>	75-76
<i>Potential Future Research</i>	76-77
<i>On the Topic</i>	
<i>References</i>	78-82

Introduction:

Daniel Dennett's work has sparked great interest in the anglo-american tradition since he first published *Consciousness and Content* in 1969 (Dennett, 1969). Since then, he has written countless articles, replies and books on topics ranging from free-will, to religion, to consciousness. In this thesis I will be looking at Dennett's negative account of consciousness, examining the role that verificationism has played in his attack of what he calls Cartesian materialism. Although verificationism has lost most of its appeal in western philosophy and has a reputation as a failed project, Dennett has informally used verificationist arguments to attack claims about consciousness that are not subject to verification. Dennett has resisted being called a verificationist and resisted his arguments being construed as verificationist. Although Dennett does coyly claim that he prefers instead to be called an "urbane verificationist", there is much hesitation from Dennett to be pigeon-holed into any kind of 'ism' especially when it comes to being associated with verificationism and instrumentalism.

But is Dennett a verificationist? Dennett does utilize a number of verificationist moves to deny the existence of qualia and if using verificationist moves is enough to be called a verificationist then he fits the bill. However, using verificationist moves might be a necessary but not sufficient condition for being a verificationist. I will argue in this thesis that Dennett is a verificationist and that his verificationism is crucial to understanding his attack on the model of consciousness that Dennett titles the Cartesian theatre. I wish to show that Dennett's own positive framework emerges from his verificationism, particularly his attack on qualia, allowing him to preserve some of our

folk psychological intuitions, while at the same time taking the scientific worldview seriously.

In Chapter 1, I will explain what verificationism is while placing it in its historical context. Since the logical positivists were deeply indebted to thinkers such as Hume, Berkeley, and Kant, I will begin by exploring the enlightenment and then trace the idea of a criterion for meaning from Hume and Kant up to the logical positivists of the 20th century. After which I will discuss how the verificationist criterion for meaning into a plurality of different views on what constitutes a meaningful statement or proposition. I will then distinguish between the different flavours of verificationism describing the distinctions between verificationism in practice and in principle and strong and weak verificationism. By describing verificationism in some detail, it is my intention to make Dennett's own usage of verificationist arguments more clear and precise.

In Chapter 2, I will be exploring Dennett's views on intentionality, folk psychology, and realism. I will argue that Dennett is committed to a form of instrumentalism, a view that he once accepted without reservation but has recently backed away from. As Dennett's intentional stance is committed to a form of instrumentalism (which is a form of verificationism), I will argue that Dennett is an anti-realist about folk psychological explanation.

In Chapter 3, I will explore Dennett's multiple drafts model as an alternative to the Cartesian Theatre model of consciousness. In this chapter my main goal will be to show what the multiple drafts model is and how it differs from the Cartesian theater. By comparing and contrasting the two respective models I hope to establish groundwork for Dennett's verificationist attack on qualia in Chapter 4.

In Chapter 4 I will discuss Dennett's eliminative materialist views on qualia and how he uses various verificationist intuition pumps to show that qualia do not exist. I will show how eliminating qualia allows Dennett to suggest an alternative to the Cartesian materialist model of consciousness so he can replace it with the multiple drafts model of consciousness (now known as fame in the brain).

Chapter 1

In this first chapter I wish to explain some of the history behind verificationism because in order to understand where Dennett is coming from with his own verificationism, we have to have some idea of where verificationism comes from. I do not wish to give an exhaustive overview of the history behind this idea since it's a complex history and certainly beyond the scope of this thesis to do so. However, I will give a brief and succinct history of the subject, so as to set some of the context necessary to understand Dennett's position. My goal is to trace the history of verificationism from the enlightenment to the 20th century. Although this will be an incomplete history, it should give us an idea of the historical basis of verificationism as well as a picture of why the verificationist criterion mattered throughout the history of philosophy. At the end of this discussion I will carve out the different types of verificationism which will give a conceptual base from which we may draw conclusions about Dennett's own verificationism.

The Enlightenment:

Verificationism has a long history in the western tradition and can be traced back to the enlightenment. During the enlightenment, there were a number of developments in

the sciences which led to a pervasive belief that everything there is to know could be understood through science. Since science was the ultimate source of knowledge it raised the question that has been around since at least Plato, where does our scientific knowledge come from? Does it come from experience? Or does it come from rational deliberation? Although it is all too easy to construe empiricism against rationalism, it would be misleading to say that there was direct conflict between the juxtaposed positions. Empiricists accepted that some knowledge arose from intuition and rationalists accepted that knowledge also arose from the senses. The differences between empiricists and rationalists are thus matters of degree and how much they weigh one source of knowledge over another.

Whereas the empiricists such as Hume, and Berkeley thought that most knowledge originates in experience; the rationalists believed that most things that could be known—could in principle—be understood by pure intellectual inquiry independent of experience. However, since Kant wrote the *Critique of Pure Reason* both views have been questioned since rationalism does not seem to yield observations which can confirm or deny propositions about reality, and empiricism alone cannot yield natural laws (that is if we accept that there is a problem of induction. If we do think that finite observations can yield natural laws, then there is no problem).

In the piece, *The Roots of Romanticism*, Berlin gives an account of the enlightenment which reduces the movement to three propositions: first, that all genuine questions can be answered. Second, all answers are knowable. And third, that all answers must be compatible with one another (Berlin, 1999, pgs.26-27). According to this view, truth about the world is like a jig-saw puzzle where propositions which are true, must fit

together in order to produce a logical, coherent view of the world. Thus, this worldview presumes that the world is rational and orderly.

For the empiricists of the enlightenment, science was the path to truth, unveiling laws of nature that reflected reality as it is. The two primary sources of knowledge for the empiricist was the mathematics and observations of science, which they believed yielded true understanding of reality. According to this view, statements about reality must be publicly communicable, testable and capable of verification or falsification otherwise they are of no utility to rational enquiry into nature (Berlin, 1977, p.163). Since science was the source of all genuine knowledge, all other forms of authority whether religious, state, or otherwise were to be looked at with a critical eye.

Furthermore, all “genuine questions” were considered to be answerable. What this proposition did not mean is that any question had an answer (Berlin, p.163, 1977). How many angels could fit on a pin would surely not count as a genuine question since it transcended possible experience and was certainly not a mathematical question. A genuine question would be ‘do objects of differing masses fall at a uniform or different rate’? This was a question Aristotle asked and gave an answer to in the third century BCE. According to Aristotle, heavier objects fell faster than lighter objects. For several centuries this answer went unquestioned. This unquestioning attitude is not what Aristotle necessarily taught us, when we want to learn about the nature of things, Aristotle deviated from his Platonic predecessors in seeing experimentation and observing nature as the source of our understanding the world around us. Aristotle would have likely endorsed, and rejoiced the next thinker who made a move that I would like to look at.

Galileo addressed the question of whether heavier objects fell faster than lighter objects, using a method that Aristotle would endorse, conducting his own investigations using a clever thought experiment, and an even more clever experiment did we get what appears to be the correct answer to the question of whether objects fall at differing rates despite differential masses or weight.

In Galileo's thought experiment, he suggests that if heavier objects fall more quickly than lighter objects (as Aristotle said), then if we tied a heavy object to a lighter object and dropped them from a building, then the heavy object should accelerate the lighter object, while the lighter object should drag on the heavier object. But this is absurd because objects cannot accelerate and decelerate at the same time. Therefore, heavier objects do not fall more quickly than lighter objects since there would be a contradiction (Brown, 1991, pgs.1-2). If Galileo's thought experiment is correct, then it leaves only one option, since objects falling at different rates leads to a contradiction, objects with different weights must fall at the same rate.

Galileo also designed an experiment where he took several spheres of various weights and measured their acceleration rates as they rolled down slides. If Aristotle was correct, then the heavier spheres should have fallen more quickly than the lighter spheres. However, Galileo found that both heavy and light spheres, rolled at the same rate of acceleration. Thus, Galileo showed empirically, that heavier objects do not fall more quickly than lighter objects, they all accelerate at the same rate.

Galileo's conclusion flew in the face of the view held by the church that everything Aristotle said was essentially correct. Why is Galileo's experiment so important? First, the experiment dislodged the Aristotelian hypothesis from the jig-saw

puzzle, showing that scientific questions have scientific answers that will reveal themselves if we use the appropriate methods of obtaining them. Second, it reinforced the notion that science was the authority people ought to look for knowledge, not the church, not the government. Instead, if we want to look for knowledge, we must seek it through the empirical sciences of physics, biology, and chemistry. Eventually people thought if we apply these methods to the study of the human subject in her political, social, and normative endeavours we could figure out the appropriate aims for private and political conduct (Berlin, 1977, p.163).

Carnap, Ayer, and many others were highly influenced by the arguments of Hume, and Kant eventually developing their own criterion of meaning based off of these thinker's views. In the next section, I will restrict my discussion to the philosophy of Hume, and Kant because they had the most profound influence on the logical positivist thinkers who came after.

David Hume and Immanuel Kant's Proto-Verificationism:

Berlin once said, "No man has influenced the history of philosophy to a deeper and more disturbing degree" than the philosopher David Hume¹. Kant had a more endearing reference to Hume who he credited with waking him from his "dogmatic slumber". Whether our view is complimentary or uncomplimentary towards Hume, it is undeniable that he had an enormous impact on anglo-american/analytic philosophy. I could not be sure, but Berlin's comment might have been in reference to Hume's project of putting philosophy on the same footing as science (Misak, 1995, p.8). The question of

¹ See page 7 of A Treatise Concerning Human Nature by David Hume in the introduction by Ernest C. Mossner.

the relationship between science and philosophy has occupied philosophical thought since the natural sciences such as physics, chemistry and biology broke off from philosophy forming their own distinct realms of inquiry. As we will see throughout the Chapter the question of the relationship between science and philosophy is one that not only occupied the thoughts of enlightenment thinkers such as Hume and Kant, but also occupied the thoughts of logical positivists as well as the later logical empiricists such as Quine.

I will begin by talking about Hume's attack on metaphysics since it had an enormous influence on Kant as well as the later positivists. In order to understand Hume's views on metaphysics it is important to understand Hume's copy theory and his distinction between impressions and ideas. Before I discuss the copy theory and the impressions/ideas distinction I want to point out a passage at the end of the *Enquiry Concerning Human Understanding* that is emblematic of Hume's views towards metaphysical thinking.

When we run over libraries, persuaded of these principles, what havoc must we make? If we take in our hand any volume; of divinity or school metaphysics, for instance; let us ask, does it contain any abstract reasoning concerning quantity or number? No. Does it contain any experimental reasoning concerning matter of fact and existence? No. Commit it then to the flames: for it can contain nothing but sophistry and illusion (Hume, "*Enquiry Concerning Human Understanding*," Section XII, Part III).

For Hume, if no abstract reasoning whether it be mathematics or logic can attest to the truth or falsity of a metaphysical claim, and if no experiment could confirm or deny a metaphysical doctrine proposed, then the question itself is meaningless and has no subsequent value to skeptical inquiry.

In the *Treatise* and *Enquiry*, Hume divides knowledge into two types: relations of ideas, which is the subject of mathematics and logic, and matters of fact, which is the

information we glean from observation and experiment. It is important to note that both sources of knowledge are for Hume arrived at from the senses via what he calls impressions. Whereas relations of ideas cannot be otherwise, and are necessary connections between ideas; matters of fact are states of affairs which could be otherwise and are therefore contingent. An example of a relation of idea would be our inability to conceive of a triangle that lacks three sides. It could not be otherwise that a triangle has three sides because of the way in which we conceptualize triangles as being three sided geometric figures that are enclosed. In contrast, an example of a matter of fact would be the hypothesis that the sun will rise tomorrow. Although it is very likely that the sun will rise tomorrow, it is only a contingent matter whether it will do so. The Earth could be pushed from the solar system by a passing rogue planet knocking it into oblivion, in which case the Sun would not rise tomorrow.

What is important to Hume about the nature of these sources of knowledge, is that if we arrive at a judgment and cannot determine that it is either a relation of ideas or a matter of fact, then we have descended into sophistry and illusion. Thus, Hume has proposed a criterion for what distinguishes knowledge from illusion by appealing to an entirely empiricist criterion justifying knowledge.

It should be noted, however, that Hume formulates this division to show that most of our mental life is constructed from impressions derived from sensory experience. If a proposition cannot be described mathematically or confirmed through some observation, then the proposition is meaningless to Hume. This is not unlike the verificationist criterion of meaning laid out by later thinkers such as Carnap, Ayer, and Schlick.

At this point, I will end my discussion on Hume by saying this: Hume is not the only thinker from the enlightenment to have an influence on the verificationism of the logical positivists (even though he is the only one I will speak at length about). Kant had a significant influence on the thinkers that would be known as logical positivists too. In fact, Ayer thought it was easy to mistake the logical positivist's project of the elimination of metaphysics to already have been accomplished by Kant. He concedes that whereas Kant's project was similar to the positivist's project to eliminate metaphysical propositions as literally meaningless pseudo-statements, Kant's task was to show that if we descend into debating metaphysical theses which transcend possible experience, we necessarily fall into contradictions (as Kant demonstrates in the Antinomies) (Ayer, p.34-35).

Yet, it's difficult not to read the Kantian project as a proto-verificationist treatise. Whereas Hume set up a criterion for meaning which involved matters of fact and relations of ideas (as discussed above), Kant sets up a criterion that allows only certain *a priori* and *a posteriori* forms of knowledge as fruitful realms of inquiry. In regard to experience, Kant thought that whenever we form judgments, the meaning concerning the content of a proposition is derived from some rule which is either confirmed or denied by sensory experience (Hanna, 2014; Strawson, 1966). Transcending these limits and descending into metaphysics cannot be an adequate grounds for science.

As we will see in the next section, a small group of philosophers in the 20th century would take the Humean and Kantian skepticism about metaphysics, as well as the principle that experience plays a fundamental role in our ability to acquire knowledge,

and run with the idea all the way to its most extreme and explicit form, particularly, the verificationist criterion of meaning.

Logical Positivism/Empiricism:

In the post-enlightenment period of the early 20th century, logical positivism/logical empiricism came into prominence in both Vienna, Berlin and in Britain as well². Influenced heavily by Wittgenstein and the new physics of relativity proposed by Einstein, the logical positivists began a project of the elimination of metaphysics—although there were other projects that logical positivists put forward—for brevity sake I will remain silent on those subjects. They sought to eliminate metaphysics with a criterion of meaning called the verificationist criterion of meaning. With the verificationist criterion the logical positivists wished to show that metaphysical propositions were not only fruitless discussions, but were literally meaningless.

The Vienna Circle:

In this section, I will first discuss the Vienna circle which was largely the center of the logical positivist/logical empiricist movement. From there, I will discuss the right and left wing of the Vienna circle and how each defended their own respective versions of the verificationist criterion of meaning, one liberalized and the other conservative. After which I will give an account of why each criterion was officially abandoned leaving the verificationist criterion an unfinished project.

² I do not distinguish between logical positivism and logical empiricism as it is challenging to draw a clear boundary between the two. I thus follow Friedman & Creath (Friedman, 1999, p.1; Creath, 2014)

The Vienna circle was comprised of scientifically trained and scientifically minded philosophers, and were one of the only philosophical movements—that I know of—whom had an official leader, namely Moritz Schlick (although one might call Sartre the leader of the existential movement, his title was not in any way official and he would have likely eschewed any such authority). Moritz Schlick—who held Mach’s chair in Philosophy of the Inductive Sciences—led the circle through weekly discussions in the problems of philosophy of science throughout the mid-twenties and thirties of the early 20th century (Uebel, 2014). The official members of the Vienna circle included: Schlick, Hans Hahn, Philip Frank, Otto Neurath, Olga Hahn-Neurath (Otto and Olga were married), Viktor Kraft, Theodor Radacovic, Gustav Bergmann, Friedrich Waismann, Herbert Fiegl, Marcel Natkin, Karl Menger, Kurt Gödel, and Rudolf Carnap. Although the circle had official members, there were also many notable visitors who would attend meetings and partake in the discussions. A. J. Ayer visited, as did Tarski, Reichenbach, Nagel (Ernest, not Thomas), Quine, and Hempel as well as others (Uebel, 2014).

The Vienna circle would meet regularly to discuss matters of scientific and philosophic interest. Among the topics that arose during their discussions was the verification criterion of meaning. What the verification criterions main function was distinguish between literally significant propositions or statements and meaningless pseudo-propositions. By eliminating meaningless propositions, the principle motivation was to show how metaphysics itself was vacuous and how science could move forward.

The Elimination of Metaphysics:

The main point of the verificationist criterion for meaning was to lay down the conditions for what propositions were literally significant and by doing so, thus eliminate

metaphysical propositions. A.J. Ayer in his *Language Truth and Logic* deals with this issue in the first Chapter that he titles, *The Elimination of Metaphysics* (Ayer, pgs.33-45). The thesis that Ayer defends is the idea that the metaphysician does not simply speculate on matters which are unknown and have no profit in their inquiry, but rather that the metaphysician produces sentences that do not have any literal meaning. Ayer's method for showing how metaphysical propositions lack literal significance is by developing a criterion for what constitutes a literally meaningful proposition, what is known as the verification criterion of meaning.

The Verification Criterion:

So far I have made it sound as though Ayer formulated the verification criterion of meaning on his own, but this is not the case as the criterion was developed by the members of the Vienna circle with the verification criterion taking on several forms before being dismantled through their own reflective scrutiny. The reason why I am focusing on Ayer's take on the issue at the moment is because I think he clarifies a number of issues such as the distinction between verifiability in practice and in principle, as well as the distinction between strict verificationism and weak verificationism, all of which are important in understanding Dennett's own arguments/intuition pumps, how the verificationist criterion(s) developed, and why it was abandoned.

What Ayer and the positivists were after was a criterion for literal meaning. What this criterion required was the conditions that would need to be set down in order for an observer to accept or reject a statement as true or false. Ayer writes,

we say that a sentence is factually significant to any given person, if, and only if he knows how to verify the proposition which it purports to express--that is, if he knows what observations would

would lead him under certain conditions, to accept the proposition as being true, or reject it as being false (Ayer, p.35).

It is important to note that Ayer is not saying that the person needs to be capable--by way of going out and making the necessary observations--of determining whether the proposition is true or false, she need only know how one would hypothetically go about determining whether the proposition is true or false or what observations or methods would matter in settling the issue. For example, if I wanted to know how many books on Carnap, the Carleton University Library owned there are several methods I could use to answer this question. I could go into the database and count how many books on Carnap the Library possessed. Despite having no interest in doing so, there is a method that I could use which would settle the question. Therefore, we can say that a specified number given whether it be 82 or 308 has an answer that is either true or false. Even if I never intend to answer the question I know what conditions would need to be in place for me to answer the question, the problem is therefore verifiable.

However, if all of the evidence that we could acquire would be consistent with the propositions confirmation or denial, then the proposition is meaningless (Ayer, p.35). For example, if a meteorologist claims that tomorrow there is a chance of sunshine, with a chance of rain, flurries, and hail, then all possible observations will conform to the prediction made by the meteorologist. There could not be any observation to demonstrate that the proposition is false since even if it was not raining, it could be sunny or hailing. As we will see in Chapter 4 Dennett attacks arguments of latter form in his attempts to undermine Stalinesque and Orwellian explanations for phi phenomena and meta contrast (Dennett, 1991, pgs.114 & p.142).

Verificationism: in Practice and in Principle.

At this point it makes sense to clarify two forms of the verificationist criterion that appear in Dennett's work, one is verificationism in practice, the other is verificationism in principle.

Verificationism in practice works like this: if we cannot confirm or deny a proposition because we lack the means whether it be technological, economical, or because of ethical concerns the proposition is said to be unverifiable in practice. If we can however, verify a proposition because we have the means, then the proposition is said to be verifiable in practice. For example, if I claim that there is a red ball in the closet behind my desk, we know exactly what procedure we would need to use to determine whether the claim is true or false. We would simply look in the closet behind my desk. The proposition is therefore verifiable in practice.

On the other hand, there are propositions that we lack the practical means whether it be economical, technological, or simply the geographic proximity to verify the proposition. An example that Ayer uses that is somewhat out of date but gets the point across. When Ayer was writing *Language, Truth, and Logic* there was no means of getting to the moon via rocket ship, since no rocket ships had been invented yet. If someone claimed there were mountains on the dark side of the moon, there would be no way in practice of verifying this proposition since the moon is tidally locked and we can only observe the luminated side of the moon. Therefore, the proposition would have been considered unverifiable in practice. That being said, since Ayer knew what would have to be the case to make the observation that would confirm or deny whether there were mountains on the dark side of the moon, the proposition is said to be verifiable in

principle (Ayer, p.36). However, if I were to make the claim that there are epiphenomenal gremlins, then we would say that the proposition is unverifiable in principle because we do not even know how we would confirm or deny whether there are epiphenomenal gremlins.

The essential difference between some proposition being verifiable in practice and in principle is this: if we have the means to test whether a proposition is true or false, right now, with little effort, then the proposition is verifiable in practice but if the proposition has some built in logical contradiction, vagueness, or impossibility of observation that prevents confirmation or falsification then the proposition is unverifiable in principle, because we do not know what would have to be the case for the proposition to be true or false. In the former case, there are many propositions that we can test the truth or falsity of but do not do so because time is precious and others have already done the work necessary to confirm or deny these propositions. Yet these propositions remain verifiable in practice because we could go out and reproduce the experimental procedures in order to reproduce the findings of the scientist. Thus, it is possible in practice for me to verify whether there is a teddy bear in the back of the red Subaru in the parking lot that my kitchen window overlooks, yet it is unlikely that I will go outside and observe whether the proposition is true or false. In the latter case, it is possible in principle to determine whether there is a star orbiting a black hole at the edge of the universe despite there being no possibility (as far as we know) of travelling to or observing an object that small at the edge of the universe. But we know what would have to be the case for us to acquire such an observation and this makes the proposition verifiable in principle. Despite our inability to acquire an observation that would confirm or deny the

proposition, we do know what would have to be the case for us to confirm or deny the proposition. Alternatively, it is impossible in principle to determine whether or not the universe is right side up or down because we do not have non-arbitrary point to decide which is right side up and which side is down.

Strong and Weak Verificationism:

The logical positivists/empiricists did not necessarily agree on how the verificationist criterion should be formulated. On the one side, there were those who were in favour of a more “conservative” criterion that argued for a strong criterion of meaning. The strong form of verificationism states that a proposition is meaningful if and only if the proposition can be conclusively confirmed or denied (Ayer, p.37). On the other side were the more liberal positivists such as Ayer who argued for the weak form of verificationist criterion which states that a proposition is meaningful only if some observation would render the proposition probable (Ayer, p.37).

Internal Criticisms of the Verificationist Criterion:

The logical positivists were not unreflective, nor uncritical of the verificationist criterion of meaning, in fact they proposed numerous formulations all of which were ultimately rejected or dismantled by logical positivists themselves.

For instance, Ayer proposed an argument against the strong verification criterion (SVC) stating that if we accept that every proposition needs to be conclusively verified, then the general propositions that we ordinarily use such as ‘apples are good for you’, “arsenic is poisonous”, “all men are mortal” are meaningless (Ayer, p.37). The reason why Ayer thinks these propositions are meaningless (if we adopt the SVC) is that he

believes a finite series of observations cannot conclusively determine the truth or falsity of general propositions (Ayer, p.37).

Positivists have bitten the bullet on the issue and have agreed that these propositions are indeed non-sense, but are an important sort of non-sense of practical utility (Ayer, p.37). This response is wholeheartedly instrumentalist in that it treats general propositions as neither true nor false but instead mere instruments of practical utility.

. The weak verification criterion instead proposed that any observation statement that was empirically relevant in regards to the truth or falsity of a proposition was a meaningful statement (Ayer, 11). As Ayer rightfully points out, “relevant” is rather vague and instead proposes that we consider observation statements to be meaningful if they are observations of an actual or possible event which would confirm or falsify a proposition (Ayer, p.11) For Ayer this was too liberal of a criterion since almost any proposition could be considered meaningful under this criterion.

As we can see from these criticisms of the verification criterion, it was never really settled what would be a non-problematic verificationist criterion. Both the liberalized and conservative verification criterion had their share of difficulties that inevitably led to their abandonment. Not only that, the verification criterion has been proposed to be unverifiable itself since it is neither a tautology, nor empirically verifiable (Taliaferro, 2003, p.454).

This last criticism is not a good one. In Euclid in order to do geometry one must accept the axioms with no further proof being possible. With the verification criterion of

meaning it can be accepted purely on a basis of its utility. By emphasizing that scientific hypotheses must in some way be testable either in practice or in principle the logical positivists have given a criterion that for all its limitations closely resembles how science is actually practiced. In this way, I think the verification criterion for all of its problems still has a place in philosophy and science even if there hasn't been a proper articulation of the criterion.

An Alternative to Verificationism:

Karl Popper in his *Refutations and Conjectures* argues that there is perhaps a broader problem when discussing confirmation as the arbiter of meaning. Since, as Hume argued, confirmations of a proposition can lead us to believe that a proposition is true or universal, when in fact, one case to the contrary can show that proposition to be false. As Hume discusses, just because the sun rose yesterday and every day before it, does not mean that the sun will rise again tomorrow. This was Popper's fear in his attempt to develop a method to demarcate scientific, from unscientific propositions. He feared that a proposition that was irrefutable would pass any test which relied on confirmation as its criterion for being scientific (or in Ayer's case, literal meaningfulness). Hence, he responded with falsification as the standard which we should hold propositions to. But this position too is problematic, since one case to the contrary does not show a proposition or a web of propositions to be false, it could merely be a problem of measurement or an extraneous variable.

Am I Arguing for Verificationism?

Although I would contend that verification is a central part of assessing conscious mental states when pursuing scientific questions, I do not think that verificationism is ultimately appealing. As I have already cited some of the problems with verificationism as a criterion for meaning, I do not think it's necessary to cite new arguments against the verificationist project. What I will say is this, I am in agreement with verificationist style arguments that Dennett makes in *Quining Qualia* (QQ) which will be discussed in chapter 4, because they demonstrate that there are contradictions in the received view on qualia.

Conclusion:

This first chapter had three goals: first, to show how the verification criterion of meaning has its origins in the enlightenment and was eventually taken up by the logical positivists in the 20th century. I discussed Hume and Kant's proto-verificationism up to the logical positivists distinction between strong and weak verificationism. Second, to show what flavours of verificationism there were, this included strong/weak verificationism and verificationism in practice and in principle. Third, I gave some of the more famous criticisms of the verificationist criterion of meaning.

Chapter 2: Intentionality, Instrumentalism, and Realism in Folk Psychology:

Introduction:

In this chapter I will be dealing with the question of whether Dennett's position on intentional states is strong enough to be a brand of realism. That is, whether or not our ascription of intentions, desires, and beliefs pick out anything real or whether they are merely useful fictions? As I will explain Dennett's position falls squarely between

Fodor's industrial strength realism, and the eliminative materialism that was once held by the Churchlands. In order to make this landscape clear I will give a brief explanation of what folk psychology is by proceeding to explain what the intentional stance is distinguishing it from the design and physical stances. I will discuss Dennett's realism about intentional states and explain how Dennett's view is a form of instrumentalism and interpretationalism. I will also explain Fodor's realism distinguishing it from the Churchland's eliminative materialism. I will then compare and contrast these views with Dennett's instrumentalism about folk psychological explanations attempting to make the weaknesses and strengths of each of these positions clear. The reason why I am talking about intentional realism is that verificationists are committed to anti-realism about anything that has no method of verification. Since Dennett does think there is a method of verification for determining intentional states through the use of the intentional stance Dennett is a sort of realist about intentional states.

What is Folk Psychology?

Folk psychology is our everyday ascriptions of intentional states to ourselves, and others behaviours. If—for example—when April says to Donatello that Leonardo is selfish, and that she does not like him, Donatello may predict that she will not attend Leonardo's birthday party. Donatello makes this inference about April's future behaviour based on the belief—expressed by April—that Leonardo is selfish, and the desire to avoid him. It could always be the case that April could show up to Leonardo's party despite him being selfish, and unlikeable, for any number of reasons, but we can predict with a fair degree of probability that she will avoid the party.

How is this possible? How can we predict what April or anyone will do for that matter? According to Dennett, there are three levels of explanation from which we can make predictions about any entities behaviour: the physical, the design, and the intentional stance. The former two explanations are not folk psychological explanations; whereas the latter is a folk psychological description of the intentional states that the entity ought to have if the entity is rational.

The three stances, and Dennett's view on folk psychology:

Crucial to understanding Dennett's views on folk psychological description is a distinction he makes between three stances that can be used to explain the behaviour of entities. The three stances include: 1) the physical stance, 2) the design stance, and 3) the intentional stance. In Dennett's philosophy of consciousness, and mind, the latter of the three is the most important for understanding his views on intentionality, and consciousness.

The Physical Stance:

The physical stance, which explains a system by the laws of physics. This strategy is reductive, highly accurate, and can describe minute details, and complex patterns that are often counter-intuitive. For example, if we want to describe how electrons will flow through a complex of circuits, this is the strategy to employ. However, if we want to describe which move a chess playing computer will make next, we would be wise to employ a different strategy. This is because using the physical stance requires powerful computational techniques that are not beyond us in principle, but are well beyond us in practice. To calculate how electricity will move through the system would require

computations that cannot be made in the time allotted between turns in a game of chess, therefore, it is better to adopt another strategy when trying to explain, and predict what the program will do next.

The Design Stance:

Another strategy is to adopt the design stance, which assumes that the system (whether it be a chess playing computer, a video game console, or a person) has a particular function that—when working correctly—should behave in accordance with the manner in which it was designed for. Dennett points out that the design stance breaks the system down into different functions, which work towards some telos (Dennett, 1978, p.4). If the system is operating correctly, then the entity should behave in accordance with its design. If the system is malfunctioning, the object may deviate from its intended design. Either way, when we adopt the design stance we are making normative assumptions about how the system ought to operate by assuming the subject is a rational agent (Zawidzki, 2007, p.35).

An example Dennett makes use of in *Freedom Evolves* is Conway's Game of Life. Although not so much a game as a simulation of a Democritean universe where very basic rules, lead to the complex behaviours of pixelated denizens that inhabit the game, which either annihilate each other, or grow, and become much more complex as the game/simulation plays on. When we look at the rules of the game, we adopt simple modus ponens rules to predict how the pixelated denizens will behave. We do not need to look at the physical level of individual pixels to know what will happen, we only need to know the rules that the “clumps” of pixels (i.e., the denizens) will follow in order to know how they will behave. By using the design stance language, we are given a

“transparent foreshortening of the tedious descriptions” of the physical level (Dennett, 2003, p.39).

One benefit of the design stance is it allows for movement from different levels of abstraction (Dennett, 1978, p.4). An example Dennett uses is an engineer following the circuits in a computer to see how the program operates. The engineer makes the assumption that the system works correctly, and in doing so, is able to reverse engineer what the chess playing computer will do next. If the engineer widens the scope, she can analyze the software in order to make predictions about the chess-playing computer without ever having to look at the hardware.

This strategy avoids the over-encumbering calculations of the physical stance, which although in principle are always available for use, in practice are limited in their utility. However, the design stance is not without its risk. What if there is some breakdown in the machinery of the chess playing computer? What if there is some error in the programming of the Game of Life? The strategy of assuming a design no longer works is the case of malfunction.

The Intentional Stance:

Both strategies are far beyond a person’s capacity to play a timed game of chess (Dennett, 1978, p.4). Instead, it is better to adopt the intentional stance, which predicts a system’s behaviour by assuming the observed system is rational by ascribing goals that it ought to have in order to achieve what it wants or desires. We also have to consider the constraints conferred on the system by the rules of the game, as well as the chess playing computer’s design. Now-a-days even the most modestly designed chess playing computer

can easily beat an experienced player with little “effort” (if set at a high enough level). But if the computer, and the subject are matched roughly evenly, and the subject adopts the intentional stance, then there is no worry about the computer mopping the floor with the subject.

This is because the chess-playing computer is roughly predictable by the intentional stance. Since we know the rules of the game of chess and know that the computer has the “goal” to win, we can treat the program of the computer as rational, with beliefs about where the pieces are on the board, and assume that it desires to protect the queen (Zawidzki, 2007, p.36). After all, if the computer’s main goal is to win, then it cannot allow a check-mate against the King, so it would do well to protect the King at all costs, which may include losing pawns, knights, and rooks in the process. At this point we can ascribe beliefs to the computer such as one or two pawns are less valuable than a knight, rook, or the queen. We can even ascribe the desire, assuming the computer to be rational, and not suicidal, to protect the King at the cost of all of these lesser denizens of the board.

We may also adopt the intentional stance when dealing with people, in fact, it would be very challenging to adopt any other stance in our everyday interactions. Imagine attempting to coordinate plans with colleagues from the physical stance. The calculations necessary to predict a group of people’s behaviour using pure physics would be astronomical, since human beings (as well as most animals) possess central nervous systems, and environments so complex that no supercomputer could predict their coordinated efforts from their atomic structures alone. Although we could accomplish this task in principle, it would be of no utility in practice.

There are also problems using the design stance to predict the behaviour of the chess playing computer. The crafty engineer can look at the software making judgments, and predictions on the basis of the system's design, but the labour, and time required to do so would be lamentable. It's one thing to reverse engineer how something is designed, it's quite another to find a partner who will allow the massive amount of time it would take to do so between turns. (This is excluding the chess tournaments that occur in the graduate student office at the Carleton philosophy department, where a friendly game may last a month or more). Although the intentional stance lacks the accuracy of the physical or design stances, as a matter of practical utility, it is best to use the intentional stance when dealing with chess playing computers or chess playing humans.

Dennett's Instrumentalist View on Folk Psychology:

In Dennett's view, just as the intentional stance leaves room for error, so too does folk psychology find itself in the position where it is not immune from error (Dennett, 1998, p.81). Although folk psychology is not immune to error, it is capable of being revised, which can be a strength. For example, science is not immune to revision, yet it is one of the most powerful tools that humanity has devised. In the same way that science is a powerful tool that is often revised due to a better understanding of nature—or for political reasons—folk psychology is often revised due to changes in the manifest image.

Dennett thinks that folk psychology is going to stick around because it's useful, in the same way that naïve physics is going to stick around despite there being much more accurate predictive models discovered by physicists. Naïve physics is our everyday conception of how physical objects will behave. If the cat walks passed the vase, and swings its tail, Jane—if she is paying attention—will tend to reach out for the vase,

because of her previous experience with objects falling, and the tendency of fragile objects such as wine glasses to break.

There are general, reflexive, and largely unconscious notions of how nature works that Jane is implicitly abiding by that are not as precise or as complex as differential calculations, but are good enough to get along in the world, when clumsy cats knock delicate objects over. First, Jane recognizes that heavy objects fall; second, she recognizes that fragile objects have a disposition to break when dropped. Since she has a goal of preserving her wine glass from being broken, she uses her folk understanding of physics to make a prediction (that she wishes to avoid) and acts in accordance to avoid the catastrophe by reaching out, and grabbing the glass midair.

Our implicit use of folk physics is so ubiquitous that we rarely need to acknowledge it explicitly in our everyday interactions with the world. It is only when we need to communicate to others that we express our folk physics explicitly. Even the physicist who knows that when Newton's apple fell from the tree, that due to gravity Newton, and the earth no less, were pulled towards the apple, and the apple towards them in a manner that defies human intuition. The relationship between the apple, and Newton is not thought of as mutually attractive according to our folk physics. Instead we appeal to what we see, and discuss the event as if the apple fell, and struck Newton without any indication that their masses are mutually attractive, rather than a one-way street.

One point that Dennett makes is a system is only considered an intentional system if it is in relation to some observer trying to observe, and predict its behaviour (Dennett, 1978, p.4). A chess-playing computer is therefore, not an intentional system, unless someone decides to play chess with it, and adopts the intentional stance towards the

system making it an intentional system. What he means by this is, only through ascribing beliefs, and desires to the system, does Dennett think it acquires beliefs, and desires. He is not saying that the system has beliefs, and desires independent of the observer. This view therefore places beliefs outside the head of the observer, and in the ascriptions of the observer who interprets the behaviour of the system.

This makes Dennett an interpretationalist about beliefs. Beliefs are not in the mind of the person making a claim about his or her beliefs, they are in the ascriptions of third person observers. By adopting this view, Dennett automatically distinguishes himself from those like Jerry Fodor, who think that beliefs are in the brain, and they are in the language of thought.

Real Patterns, the Intentional Stance, and Interpretationalism:

Philosophers of science distinguish between scientific realism, and anti-realism concerning entities. Where scientific realists agree with anti-realists is on the ontological status of observable entities. For example, both realists, and anti-realists will agree that planets, fossils, and stars exist. Where the realists, and anti-realists disagree is when it comes to unobservable entities such as electrons, protons, and other particles that cannot be directly observed. Whereas the scientific realist will contend that the electrons, protons, and other unobservable particles are real; the anti-realist argues that they are merely useful fictions (Okasha, 2002, pgs.58-61).

For Dennett, beliefs are real patterns that are to be found in behaviour. In *Three Kinds of Intentional Psychology*, Dennett claims that beliefs are *abstracta* much like centers of gravity or the equator (they are mathematical or logical constructs) (Dennett,

1987, p.53). It is important to note that Dennett is not saying that beliefs are real in the same sense that a brick, a city, or a book is real. We don't stub our toes on the equator but it does not mean the equator is not a real objectively definable coordinate that can be drawn. Dennett is dealing with scientific abstracta which are abstract objects that are somehow useful for scientific theories. Hence, Dennett is a form of instrumentalist about beliefs.

We must then ask, what makes a scientific theory of belief attribution useful? Well, for Dennett, a scientific theory is useful on pragmatic standards such as whether or not the theory can predict, and adequately explain a pattern. For example, say Jones says he believes he likes Jane. But Jones also says that he likes Jaimie. How do we choose between the two beliefs? The answer—for Dennett at least—lies in practical considerations, and would be up to the observer (Dennett, 1998, p.118). If Jones has made both claims more than once, then Jane has observed a pattern, because it (the verbal expressions to be interpreted as beliefs) can be picked out more simply than a bit map (and that's all there is to a pattern for Dennett). Since all she needs is an idiosyncratic consideration, then Jane may choose to think that Jones likes them both or Jones only likes either Jane or Jaimie, and it would be entirely up to her as the interpreter to attribute the belief (whichever way she chooses) to Jones or not.

To What Extent is Dennett Plausibly a Realist? Or Not?

Dennett claims that since he thinks that beliefs are an objective phenomenon it makes him a realist about beliefs (Dennett, 1987, p.15). But he thinks that beliefs can only be discerned by adopting the intentional stance which makes him an interpretationist (Dennett, 1987, p.15). For Dennett to be a realist about beliefs he has to accept that

beliefs are more than patterns of behavior discerned from the intentional stance, he must think beliefs are features of neurological structures in the brain, otherwise beliefs would be mere useful fictions that we infer in order to predict behaviour. That would make him an anti-realist about beliefs, not a realist, specifically an instrumentalist. In order for Dennett to be a realist about beliefs he would have to think that beliefs can exist independently of interpretation from an external observer. The only way for Dennett to do so would be to identify beliefs with neurological activity (which is plausible) or to adopt some sort of dualist picture of the mind where beliefs exist in some other realm (the latter of which is unlikely). Dennett likens the former as a sort of realism about beliefs where the person having the belief is likened to being infected with a sort of “virus”, an objective fact about the internal state of the subject (Dennett, 1987, p.14).

Why does Dennett consider himself a realist about beliefs? Dennett thinks that beliefs are an objective phenomenon and that is what makes him a realist about beliefs (Dennett, 1987, p.15). What is interesting about Dennett’s view is to be a “true believer” is merely to be “reliably predictable via the intentional strategy” (Dennett, 1987, p.15). As we saw earlier, the intentional strategy works by establishing the context (including its evolutionary and learning history), what beliefs the system ought to have given its present situation (is the object edible or poisonous?), the information the system has access to (the object is a rock) and if the system is predictable using the intentional strategies assumption of rationality (do not eat rocks). It is these specific conditions that make the person (or system) an intentional system that has beliefs and desires.

But can a belief exist independently of an interpreter who adopts the intentional stance? Or more pointedly, are beliefs mental states that are manifestations of the brain

(or any structure that functions in a similar manner)? If so, then Dennett can rightfully be called a materialist, and also a realist about beliefs such as mental states that are stored and put to use by at least some organisms in the production of action some of the time. This view would preserve our folk psychological intuition that beliefs are causally efficacious insofar as they provide causes for action, while at the same time resist the temptation to eliminate belief as a fundamental process of human thought and cognition. Of course beliefs can still be ascribed (and may only be intelligible) using the intentional stance, however, they would only be part of the story which could in principle be told. Under this view a brain might—in principle—be read using brain scanning technologies in order to determine whether there was a fact of the matter concerning whether someone had a belief or not, but that technology may be in practice, forever beyond our ken. Thus, the position could maintain epistemic neutrality in regard to our ability to know about beliefs from the perspective of neuroscience, but at the same time accept a materialist position that the brain is the vehicle of beliefs and the target of interpretation.

The latter view we can call neurological realism. If Dennett is committed to the former view that beliefs are only ascriptions we make in order to explain people's behaviour because they have predictive value, it's hardly that strong of a realist position. Dennett's interpretationism is analogous to the thinking underlying quantum mechanics, Putnam gives a very good analysis of the distinction between quantum mechanical thinking (at least outlined by Bohr) and classical mechanical thinking. Putnam writes,

The theory... applies to the system—say a system of elementary particles, or a system of fields and particles. As in classical physics, the system can be quite small—one or two or three particles—or it can “in principle” be quite large. But—here is the curious feature which was not present in classical physics—any application of theory requires that, in addition to the “system” being talked about, there be “apparatus” or an “observer” which is not included in the system. In principle, then there is no “quantum mechanical theory of the whole universe” (Putnam, p.4)

The final point Putnam makes is unimportant for our purposes, but his point that quantum mechanics—under—Bohr’s Copenhagen interpretation—is observer-dependent is revealing to our discussion of Dennett’s interpretationism. Like Dennett, Bohr (and colleagues) think that the apparatus and the observer using the apparatus, although not part of the system being observed, must be applied when formulating the theory of the system, otherwise there cannot be any meaningful description of the system in question. Both Bohr and Dennett might agree that the particles in the system (or neuronal networks in the case of beliefs) exist independently of observation, they also agree that a theory of a system requires an observer and similarly they both talk about the observer adopting or (in Bohr’s case, described) by a particular stance. In Bohr’s quantum theory, the observer must be treated like classical mechanical system (Putnam, p.4), in Dennett’s view, the observer can only describe the system if she acts as an interpreter applying folk psychological, mentalistic talk about the person in general by ascribing beliefs that the subject “ought” to have given the context she finds herself in.

Are beliefs in the head for Dennett? One answer has been, no. Patricia Kitcher captures this interpretation of Dennett’s view when she writes,

His [Dennett’s] considered view about contentful beliefs sits right on the knife’s edge: people truly have such beliefs, for we get true behavioral predictions on this assumption; yet, nothing inside a person’s head is a belief (Kitcher, 1990, p.127).

There is reason to read Dennett this way. After all, Dennett does not attempt to map beliefs onto neurological structures. This move (or lack thereof) is important for it allows Dennett to remain neutral about the machinery underlying the intentional states we are interested in (Zawidzki, 2015, p.597). The brain is not at the level of discourse that the intentional stance targets, it’s instead concerned with behavioural patterns of whole

persons. As we saw earlier, for Dennett, a person can only be ascribed beliefs from the perspective of the intentional stance which depends on assumptions of rationality given what the person ought to believe given its circumstance in order to make accurate predictions about her behaviour (Shoemaker, 1990, p.212).

If this reading is correct, and beliefs are not located in the brain, then how can Dennett be a realist about beliefs? In my view he cannot be. Dennett can be a realist about patterns of behaviour, and about the objectivity of the interpretations made from the intentional stance; what he cannot be is a realist about the status of beliefs if they do not correspond to patterns in the brain which sometimes cause action. If Jones walks into the kitchen on autopilot and opens the fridge door out of habit, we cannot ascribe the belief to Jones that he was hungry, even if he habitually behaved like this on a regular basis thereby producing a real pattern. Ascribing any belief to Jones in such a state would miss an important feature of our concept of belief, that beliefs are causally efficacious thoughts that can either accompany a behaviour or not. Conversely—as the former example suggests—behaviours can occur without any beliefs whatsoever.

To be clear, I am not claiming that there is some central location where beliefs happen, where consciousness is centrally located, but rather that beliefs are constructed by various subsystems in the brain. In this sense, I am agreeing with Dennett that the brain is a massive parallel processor, but because I agree with him that multiple subsystems underlie consciousness, I can also claim that it is also possible for there to be unconsciously activated behaviour that are nothing like beliefs.

Shoemaker and Lycan both read Dennett as an anti-realist about beliefs

(Shoemaker, 1990, 212; Lycan, 2003, p.191). Both Shoemaker and Lycan read Dennett as rejecting the idea that beliefs are inner causes. For example, Lycan writes that,

Dennett... has been particular concerned to deny that beliefs and desires are causally active inner states of people” and instead likens our ascriptions of beliefs and desires to “calculating devices” (Lycan, 2003, p.191).

I would like to say a few words about Dennett’s commitment to one central tenet of functionalism, namely that beliefs are multiply realizable because I think it makes good sense of the reason Dennett does not think that beliefs can be identified with specifiable structures in the brain. A primary example of Dennett’s view that beliefs are multiply realizable is the thought experiment that could be called Murder in Trafalgar Square (Dennett, 1987, p.54-57). In the thought experiment Dennett describes an event where Jacques shoots his Uncle who subsequently dies. It is clear that Jacques believes that his Uncle is dead, since he was the murderer who was the participant and witness of his Uncle’s demise. Sherlock, Tom, and Boris also share the belief that a Frenchman murdered a man in Trafalgar square, yet none of them learned about the incident in the same way. Sherlock knows that the Frenchman murdered the man because he was the detective who apprehended Jacques, Tom reads about it in the *Guardian*, while Boris reads about it in *Pravda* (presumably before Boris sold it) (Dennett, 1987, p.54-55). Although they share the same belief, Dennett denies that the concrete instantiation of the belief could be pinned down in the same localized area of each brain of each individual.

More interestingly beliefs in Dennett’s view do not exist in isolation as individuated structures, beliefs are connected to an indefinite number of other beliefs which have effects on behaviour such as verbal utterances. We can call this view the no-

Robinson-Crusoe-belief-hypothesis. As Dennett writes, “judgements, unlike beliefs, happen one at a time” so Dennett thinks that beliefs do not happen one at a time, they happen in concurrence or in a web of intermixed beliefs. Dennett continues, “we have an indefinitely many beliefs, but can only have one thought at a time” (Dennett, 1978, p.48). To exemplify how beliefs do not occur in isolation of other beliefs, we can examine one of Dennett’s intuition pumps in *Brainstorms* (Dennett, 1978, pgs.44-45). Dennett asks us to imagine a neurocryptographer (a brain coder) who inserts—through some unexplained means—a belief into the brain of a subject named Tom the notion that he has a brother in Cleveland. When sitting at a bar, a patron asks him whether he has “any brothers or sisters?” Tom says “I have a brother in Cleveland”. When the friend asks Tom what his brother’s name is he will—according to Dennett—have one of two responses: 1) I don’t have a brother in Cleveland or 2) I don’t know his name (in recent talks Dennett has also added that Tom could start confabulating).

The question Dennett is asking is has the neurocryptographer successfully inserted a new belief in Tom? Dennett’s answer is ‘no, the cryptographer has been unsuccessful’. What Dennett has in mind here is that Robinson-Crusoe-beliefs cannot exist. If a cryptographer aims to write a belief into the brain of Tom, he must write in an indefinite number of other pieces of information into poor Tom’s head otherwise he will be unable to respond in any way that reasonably resembles having a belief about having a brother in Cleveland. In other words, even if it is in principle possible to insert individuated sentences into the brains of people, it is unlikely that neuroscientists could ever write a belief into the brain of an individual without inserting enumerable other thoughts that cohere with one another, which may be forever beyond possibility.

There's good reason to suppose that the conclusions from the neurocryptography examples, that Dennett thinks it unlikely that beliefs are stored in the brain in a homogenous functional location that is species wide. But it's still possible that Dennett thinks beliefs are constructed by parallel processes in the brain which is precisely what Dennett's multiple drafts model predicts about conscious narratives. If Dennett is saying that beliefs are constructed by the brain, but not accomplished by a unique structure dedicated to beliefs, then Dennett is still a realist about beliefs, because he's committed to the existence of parallel structures processing away that produce beliefs.

If this was Dennett's view on beliefs I could certainly agree that his view was robust enough to be rightfully called a realist position. Beliefs would just be the active processes that occur in a massively parallel processor. We may never be able to use a cerebroscope to read the beliefs from someone's brain, but beliefs would still be real regardless of us being able to read them. But if Dennett thinks that beliefs are only an indispensable way of describing and predicting other people's patterns of behaviour, then Dennett's "realism" is quite weak.

Eliminative Materialism & Industrial Strength Realism:

Some thinkers believe that folk psychology is so radically false that it will eventually be displaced, and eliminated by a complete neuroscience (both Churchlands). This view known as eliminative materialism sees folk psychology on par with antiquated theories in science (such as Aristotle's description of how objects of differing weights

fall), and predicts that it will be subsequently replaced by a better theory or description (such as Galileo's description via experiment). The Churchlands think that appeals to beliefs, and other intentional states will inevitably be replaced by the language of neuroscience (Ramsey, 2013). There are good reasons to be skeptical of this view as we will see later. Unlike the Churchlands that deny that folk psychological ascriptions of belief have any role in a complete science of the mind, Fodor proposes what Dennett calls industrial strength realism, where beliefs are physical structures in the brain. According to this view, beliefs are not dispositions to behave in ways that are predictable from the intentional stance, they are neurophysiological structures in the brain (Murat, 2015). These neurophysiological structures produce mental representations which due to their relationships between one another, create the syntactic, and semantic structures that generate thought (Murat, 2015). This underlying language, which supposedly constitutes all thought, is called the language of thought. The language of thought is a lot like the binary code which programs are written on. Imagine a computer system which has a word program on it. If the binary code is the language of thought, then the English sentences in the program are what the language of thought is translated into. The language of thought (LOT) is supposed to be basic symbols, and relationships between symbols which any natural language are translated into (Brook & Stainton, 2000, pgs.54-55).

One of the virtues of the LOT is that it explains how a Japanese person can have the same thought as an English person, but in a different language. If the LOT expresses propositions which can be translated into any language, it supports the notion that thought is separate from language. It seems intuitive that when we have a thought, that the

language we use to express that language is separate from the thought itself. For example, whenever we imagine a word to express a thought, it seems intuitive that a thought is different from the language we use to express that thought. If the LOT hypothesis is correct, then there is an explanation for at least some instances where we search for words to express our thoughts.

One problem that the LOT hypothesis encounters is that it invokes an additional system of language below the languages that we already know exist. If the LOT exists at a level below natural languages such as French, and Chinese, then it requires an explanation as to how the LOT came into existence in the first place. In the same way Dennett objects to the Cartesian theater based on an objection that the Cartesian materialism requires a homunculus, the LOT—as an explanation for language—could be thought of as invoking an additional homunculus where there is no need for one.

Paul, and Patricia Churchland have attacked folk psychology in several ways, one argument that they present is against folk psychology as a theoretical enterprise. As the Churchlands have argued, folk psychology has no explanatory or predictive force when it comes to explaining the nature of memory, learning, consciousness, and even mental illness. Folk psychology is generally silent on these issues, but neuroscience has a lot to say about these issues. Take memory, and learning as examples, because of examples from patients, and animal models, we know that the hippocampus has a crucial role to play in memory consolidation. If a person has their hippocampus destroyed, then they lose the ability to form new memories (anterograde amnesia). Neuroscientists have also observed structures such as the frontal lobe which are crucial for attention, inhibition, and higher reasoning skills which when damaged can cause diminishment in these capacities.

Other structures have been associated with spatial recognition, the ascription of the 'I' to the self, and areas responsible for language processing.

As the various areas of the brain are mapped out in terms of their functions, it is challenging not to see folk psychology (when thought of as a theory) as a poor contender next to the various advancements in neuroscience. This is true, but as Dennett notes, folk psychology is remarkably accurate, and easy to use (Dennett, 1998, p.82), and since neuroscience, whether it be complete or incomplete, is difficult for those of us without proper neuroscientific training to wield, folk psychology will be indispensable for making accurate predictions about other people's behaviour. Since most people have concerns that require no knowledge of neuroscience (complete or incomplete) to be able to predict the behaviour of others, from a purely practical standpoint, it seems unlikely that neuroscience will eliminate or displace folk psychological explanations, and folk psychological understanding of behaviour.

Folk Psychology as a Theory:

One crucial point that the Churchlands make in regard to folk psychological explanation is that it is a theory. If folk psychology is a theory, then—like any scientific theory—it can be replaced or eliminated by a better theory (a completed neuroscience). Dennett takes issue with folk psychology as a theory. But what makes a theory a theory for Dennett? Or better yet, what makes a theory any good?

First, a theory to be any good it must account for all of a particular phenomenon observed hitherto, as well as make novel predictions about the future (Dennett, 1996, p.158). General relativity easily fits the bill, since it accounted for all of the observations

that Newton's theory explained, as well as observed phenomena that could not be explained by Newtonian mechanics. Einstein's theory also predicted phenomena that had not previously been observed. If folk psychology is a theory, it would have to account for all of the observed phenomena we know about now (this includes priming, primacy and recency effects, as well as mental disorders). However, folk psychology has nothing to say about priming or primacy, and recency effects or mental disorders. If folk psychology is a theory, then it cannot be the correct or complete theory of human behaviour if it can only predict behaviour most of the time.

Novel predictions often play an important role in what makes a useful scientific theory. A scientific theory that does not make predictions that explain phenomena that we would like to know more about, or that we had not thought of as being previously possible, certainly decreases the attractiveness of the theory, since it does not expand our knowledge in any way. Take for example the center of Dennett's lost socks drawer. Although we could generate predictions about where the center of lost socks will be at a given time, the predictions would lack novelty, and utility in increasing what we know about the world. In short, it wouldn't make a difference if we made accurate predictions about where Dennett's lost sock center happened to be. A novel prediction about the way photons will behave when entering a hypothetical gravity well is something we would be interested in.

The force of a novel prediction is not only in teaching us about things we hadn't previously considered, but it gives some reason for thinking the theory is correct. Although no one would say that verification of a novel prediction is proof that the theory is true, it certainly motivates us to take the theory seriously.

Is Folk Psychology a Theory?

Although folk psychology makes novel predictions, they are vulnerable to error. Consider for example how vulnerable we are to cognitive biases, heuristics, and logical fallacies that can often lead us to erroneous predictions about our own, and other's behaviour. For example, one belief that people generally have is that in situations where there are many people, a person who is in distress will be more likely to be helped by a bystander than in situations where there are few or one bystanders passing by. However, this is not the case. People in distress are more likely to receive help from a bystander in situations where there are fewer people around (Santrock, et al, 2005, pgs. 659-660). Another example is how people often seek out evidence that confirms their beliefs rather than evidence that seeks to disconfirm their beliefs. Furthermore, there are examples where people believe what the majority of people believe (*argumentum ad populum*). Another example where folk psychology fails is when people say they suffer from a poor memory. As psychology has shown, there appear to be several different memory systems. There is short term memory, long term memory, episodic memory, and semantic memory.

But fallibility is not reason to discount a theory, as a theory. There are many theories that have been shown to be false, yet were nonetheless theories. Moreover, just because folk psychology sometimes errors and cannot explain everything, it does not exclude it from theory status. There is no unified theory of the mind as of yet, but there are many theories that explain pieces of the puzzle. There are also theories that work well in some areas, but do not work well in others.

Dennett makes another point against eliminative materialism that I think should be raised. There is good reason to believe that folk psychology is not a theory. After all, theories are usually fairly explicit in their descriptions of propositions, and laws, but in folk psychology most of the attempts to describe the world never get formulated as anything like laws or propositions (Dennett, 1998, p.82).

Although I agree with Dennett that folk psychology is not a theory, I have to disagree with Dennett's view that folk psychological explanation does not get explicitly cited in terms of theorems or laws, nor do I think that the lack of explicit theorems or laws, immediately disqualifies folk psychology as a theory. For example, we can imagine Jones saying to Jane "be careful. Fang is a mean dog, if you pet him he will bite" (Cain, 2002). What is being implicitly appealed to here is a law, which states that 'all mean dogs bite'. The law that's being appealed to does not need to exist, after all, there could still be what we would call mean dogs that do not bite, but for some other reason are mean, instead, the point is that the propositions of folk psychology can easily be thought of in law-like forms. These laws may be wrong, but in principle, there is no reason the propositions of folk psychology could not be translated into psychological laws, even if they happen to be incorrect.

As for folk psychology lacking theorems, I agree with Dennett that people's usage of folk psychology generally does not involve the use of theorems for predicting or describing the behaviour of others. However, I do not think that the use of theorems is necessarily a facet of what makes a theory a theory. For example, natural selection—although it does involve theorems at some scales of discourse—it as a theory, does not have an overarching theorem to describe everything. Nor do behavioural theories of

learning necessarily need to be understood in terms of theorems. For example, when looking at different reward schedules, there need not be any mathematical theorem that describes the exact ratio of reinforcement. The basic idea is simply that behaviours that are reinforced tend to persist, those that are punished tend to extinguish, and removal of an aversive stimulus when positive behaviours are demonstrated. The basic theory may require mathematical theorems to make more specific predictions, and descriptions, but the fundamental theory behind behavioural learning theories does not require a mathematical theorem.

It seems that there is nothing essential about mathematical theorems, and that folk psychological explanations can take the form of psychological laws, even if they are false. But what reason is there for not thinking it's a theory?

One view is that what makes a theory a scientific theory is that it is falsifiable. This view was proposed by Karl Popper in the 20th century. Popper thought that using confirmation as a criterion for what constitutes a theory cannot be right, since it's very easy to find numerous confirmations for almost any theory regardless of whether it is true or not. Instead, Popper proposes—and this is contentious—that one observation to the contrary of what the theory predicts proves that the theory is false.

How does falsification result in a scientific theory? What Popper thought was that by eliminating theories that could be proven false, and by ignoring theories that could not be proven false, we get closer to the truth by bringing forth theories that have been tested rigorously. Hence, Popper has in mind an evolutionary perspective on scientific development. Those theories that are fit enough to survive rigorous testing are true, while those that do not are false or not properly scientific to begin with.

There are problems with this view of course. In the paper *Two Dogmas of Empiricism*, Quine makes the point that a theory cannot be tested as an individual prediction isolated from other propositions. Drawing from Duhem, Quine developed the idea that no proposition could be isolated, and tested, away from the network of propositions which make up the theory. Any theory is going to have any number of auxiliary hypotheses, which generate any number of predictions, all of which are revisable in principle (Searle, 2003, pgs. 6-7). If this network of propositions does not produce any atomistic propositions that can be isolated for falsification of the whole, then this leaves an issue for Popper's theory. The issue being that Popper's theory depends on propositions being capable of being isolated, and falsified, therefore proving the entire theory false.

Folk psychology also makes numerous predictions that if written out would encompass an almost infinite number of propositions of which numerous other hypotheses could potentially be generated. If we prove that one prediction is false, or find one observation that disconfirms the proposition in question, it still leaves the remainder of folk psychological explanations untouched. This is why Dennett is correct in saying that folk psychology is revisable. Like any scientific theory, if folk psychology is a set of interconnected propositions, then any folk psychological explanation that is shown to be false, will not damage the theory as a whole, as long as we are prepared to revise our folk psychological explanations in response to new information.

The reason that I think that folk psychology is not a theory, is not because it lacks theorems or laws that are used to explain people's behaviour. Nor do I think it's because it lacks predictive precision. Instead, I think folk psychology is not a theory because it is

far too broad and unstable. Theories—although malleable to some degree—have a structure that is somewhat stable. Since beliefs vary amongst different people based on economic, religious, philosophical, political and other backgrounds that influence not only what we believe, but how we come to believe and predict others behaviour, I would contend that folk psychology is what we want to explain with cognitive science, experimental psychology, neuroscience, etc, and as such is not a theory at all.

Where Do Dennett and I Agree and Disagree?

Where I agree with Dennett (and almost everyone else) is that ascribing intentional states to others has predictive value. Dennett and I also agree that folk psychology is not a theory, but for different reasons. Where Dennett, and I disagree is on the status of beliefs. Beliefs, desires, intentions, are all in the head. Although behaviour, and action are at least sometimes manifestations of beliefs, and beliefs do have causal roles to play in behaviour at least sometimes, it's not the case that beliefs are merely ascriptions we use to describe, and explain behaviour from the intentional stance. They are not *abstracta*, they are *concreta*, that exist in the brain, and they are just as real as the metal in dollars. Although beliefs may not be able to be mapped in the same way that neuroscientists can map other functions, it does not mean that beliefs are not real in the same sense as the metal in dollars.

Conclusion:

In this chapter I explained the differences between the three stances: the design stance, the physical stance, and the intentional stance. I then explained how Dennett's instrumentalism about scientific theory works, as well as Dennett's interpretationism

about intentional states. After which, I compared, and contrasted the Dennett's mild realism with Fodor's industrial strength realism, and Churchland's eliminative materialism. I then argued on the side of Dennett and against Churchland that folk psychology is a theory, and sided with Fodor that beliefs are in the brain, but left it an open question whether beliefs take the form of mentalese.

Chapter 3: Two Models of Consciousness: Cartesian Theatre and Dennett's

Multiple Drafts Model/Fame in the Brain

In this third chapter I will explain Dennett's model of consciousness (the multiple drafts model/fame in the brain) as well as his main target known as the Cartesian theatre. The reason why I wish to discuss the Cartesian theater against the multiple drafts model is that Dennett's verificationist arguments that he has presented over the years against qualia, phi phenomenon, and metacontrast (which will be examined in the next chapter) can be seen as a progressive development towards elimination of the Cartesian theater making room for Dennett's own model of consciousness. First, I isolate Cartesian dualism (CD) explaining how it evolved into Cartesian materialism (CM). Second, I reconstruct Dennett's model of consciousness. Third, I defend the multiple drafts model as a viable alternative to CM. The goal of this chapter is to connect Dennett's verificationist attacks on qualia, the Mary the superscientist thought experiment, as well as genuine experiments such as phi phenomenon and metacontrast so that I may argue that Dennett's replacement for CM is a genuine starting point for a theory of consciousness.

Cartesian Materialism: A Model of Consciousness:

Descartes was one of (if not) the most influential philosopher of the 17th century producing major advances in the areas of mathematics, and natural science. Like Hume, Descartes wanted to apply the same techniques that had been so successful in the natural sciences to the human subject. A source of inspiration for Descartes was the gardens he would take leisurely walks through which had moving, human-like machines that would move, and interact with one another as if they were having conversations and participating in normal human affairs³. What astonished Descartes about these machines was that they were automatons lacking anything like consciousness. The behaviour of the automatons could be entirely explained by a series of pressure plates attached to gears and levers which when stepped on would cause the machines to act as though they were people with thoughts and feelings (Flanagan, 1991, p.2).

Descartes thus began thinking about the body of animals as machine-like, and theorized that they could be understood under the same basic principles as physics (Hadfield, 2001, p.3). One important result of his theorizing was that he came up with the reflex-arc⁴. Although Descartes overstated the ability of the reflex to explain animal behaviour (and at least some human behaviour) Descartes had stumbled upon a new way of thinking about the body, the body could be thought of as a machine.

One troubling feature of the mechanistic view of the body is the placement of the mind. If indeed the body is a machine (or machine-like), then it raises the issue of how executive control is possible? How could a mere machine be able to make free choices?

³ The French Royal gardens were the 17th century version of what Flanagan calls Disneyland.

⁴ The reflex-arc is still widely used today as an account of how certain behaviours that occur in both humans and non-human animals when an unconditioned stimulus causes unconditioned response that is automatic, and unconscious.

To answer this question Descartes invoked the notion of an immaterial soul which interacted with the physical mechanisms of the body providing a mechanism for executive function of the body machine which would otherwise be reflex driven (Flanagan, 1991, p.23).

Discarding Cartesian Dualism:

Dennett does not give any credence to the possibility of substance dualism endorsed by Descartes and spends a little time in CE problematizing the issue. In Descartes view the body simply transmits information it receives from the eyes, the ears, and taste buds to the pineal gland where it is then processed by the immaterial mind and transmits the information back to the body carrying instructions for how to act accordingly. Thus, if Robinson Crusoe touches fire there are three things that occur: 1) the heat from the fire causes damage to the body which is transmitted through the afferent nerve cells up to the brain, 2) the pineal gland fires off a message to the immaterial mind, 3) the immaterial mind processes that information, and projects the decision back to the pineal gland where the information then travels down the efferent nerve cells to produce action (Dennett, 1991, pgs.34-35).

There are a number of objections against the idea of Cartesian interactionism: 1) if indeed there is an immaterial mind that acts as a receiver and executive dispenser of information that poises the body for action, this raises the issue of how a non-material mind can interact with a material body (and vice versa)? 2) For me the strongest objection to dualism is there is no evidence for it and why should there be? Even if it were true that the mind was immaterial, there would be no way of testing it. In which case it could make no difference if the mind was immaterial or not. 3) A third objection is

not so much an argument as a rule of thumb, if the mind can be explained with simpler mechanisms that yield novel predictions about how the mind works, then why invoke an entity (an immaterial mind) that only invokes greater complications to our understanding of the mind. 4) This final objection relates to three, the interactionist theory is incapable of producing novel predictions. Therefore, why should we invoke an immaterial mind that makes no predictions, raises more questions about causal interaction than can be answered such as how such an interaction would be possible between material and immaterial structures in a subject? Furthermore, why should we accept that there is an immaterial mind if the mind/brain can be explained parsimoniously without appeal to exotic, ill-defined properties that do not have any measurable effects in the world? And finally, why should we accept a theory that produces no measurable predictions? For these reasons (as well as others) we should be cautious about accepting substance dualism.

The Cartesian Theatre & Cartesian Materialism:

There are a number of good reasons for rejecting substance dualism, but what about what Dennett calls Cartesian materialism? Cartesian materialism can be described as accepting materialism, rejecting Descartes dualism/philosophy of mind, but forgetting to eliminate the central place where sensory input funnels into a central location in the brain (Dennett, 1991, p.107). What Cartesian materialism (CM) predicts is there is a firm dividing line between conscious and unconscious mental states. Thus, there are facts about where consciousness begins and ends. As we will see in the discussion about phi phenomenon and metacontrast illusions Dennett thinks there are no facts of the matter concerning when consciousness begins and ends. But before I discuss Dennett's

verificationist rejection of CM, I must first introduce his own model of consciousness, what is known as the multiple drafts model (MDM).

Multiple Drafts: A Model of Consciousness:

In this section I will explain the central tenets of MDM are in order to show what Dennett's replacement for the Cartesian theater looks like. By my count there are four essential propositions which make up MDM. All of which demand some elaboration.

- 1) All mental and varieties of thought manifest through parallel processes of "interpretation and elaboration" from sensory inputs. Information that is being processed in parallel is always subject to "editorial revision" (Dennett, 1991, p.111).

Dennett thinks the brain is a massive parallel processor which receives input, and transmits information throughout the brain where the information goes through a variety of interpretation channels and is elaborated on. This process can be analogized to recording an album. When a band is recording, if they do not have much studio time (since studio time is expensive) thrifty musicians can record multiple instruments at once, saving time and therefore money. The studio will be set up with a number of mics in front of the drums, there will be a mic for the singer and other instruments such as guitars will be plugged into the computer system.

When the band starts playing, the information travels into the computer from the various recording apparatuses where it can be mixed on multiple tracks and layered where volumes can be adjusted, gain can be added to guitars, and other effects can be added and elaborated on producing what we would call a song. If the singer sings slightly flat on a note, the recording engineer can always edit the track bending the note upwards

tuning the information so that it produces the appropriate note. For example, when Jimi Hendrix recorded *All Along the Watchtower* he sat in the studio continuously revising his guitar solo experimenting with different objects to produce the slides in the riffs (he eventually settled on a zippo lighter). Whenever we listen to the radio version *All Along the Watchtower* we hear the solo that was finally selected to be heard by the general public, but that is not the only version of that solo, nor the song.

The process of recording a song on multiple channels is not unlike what Dennett has in mind when he talks about how the brain processes information. For example, the McGurk effect is an illusion that occurs when auditory information is paired with visual information which alters how the auditory information is interpreted by the subject. One example of the McGurk effect from a *Horizon* documentary called *Seeing is Believing* (not to be confused with the Dennett paper under a similar title) (Dennett, 1996, p.158) shows a man saying “ba, ba, ba,” over and over again with his lips making the puckering lip movements commonly associated with a “ba” sound. Then they flip to another scene where instead of making the puckering lip movement, the gentleman switches to a “fa” movement which involves biting down on the lower lip and releasing the sound from one’s mouth. We as individuals then hear “fa” even though the audio-track remains the same. The upshot is that how we perceive visual information can affect what we hear. Hence, audio information is subject to revision on the basis of conflicting visual information.

What is common to both examples is the emphasis on parallel processing of information and the tendency towards edits that can continue on indefinitely. In the first example, there were multiple tracks for each instrument which when put together make

the appearance of a single song. In the second example, there are two stimuli: 1) audio and 2) visual. Since the visual stimulus affects the way the auditory is perceived it shows that experience can be edited depending upon what information is being processed at a give moment.

- 2) Feature detections only need to be made once, they do not need to be made again at some master discriminator.

Dennett is making two claims: one positive and the other negative. The positive claim is that discriminations of some feature need to only be made once. The negative claim is that there is no representation (re-representation) to a centralized place in the brain where everything comes together.

The second claim is clear enough, there is no metaphorical or literal homunculus in the brain that demarcates between conscious and unconscious processing. This element of MDM is in place to avoid the Rylean problem that arises when one explains a complex phenomenon with an equally complex explanation sharing the features of that which it intends to explain (Dawson, 2013, pgs.46-47). For example, by invoking an equally complex structure (a homunculus) to explain an equally complex phenomenon (in this case, consciousness) an infinite regress occurs where each level of explanation requires an equally conscious structure to explain the higher level of consciousness.

The first claim is positive in the sense that it makes a prediction that feature detections only need to be made once. What Dennett means here is that whenever a discrimination is made, for example, between orange and red, the information does not have to travel to an inner homunculus to be represented (Dennett, 1991, p.113).

- 3) Fourth, these processes of spatially and temporally distributed discriminations produce something like a narrative that continues indefinitely into the future (Dennett, 1991, p.135).
- 4) In CE Dennett gives us a fifth tenet. This tenet is what he calls a probe that interacts with the stream content fixations producing different narratives, at different points, producing different effects.

The probe has numerous problems and questions associated with it, specifically, what sort of mechanism could accomplish such a task of identifying and selecting which content fixations become conscious? Wouldn't this probe have all of the problems associated with a homunculus? Such as explaining x with y which seems have all of the abilities of x? Furthermore, isn't there a problem of infinite regress that precipitates as a consequence of having a probe which seems very much like a homuncular explanation? Dennett later replaced the unfortunate probe and the multiple drafts metaphor with fame in the brain. As Dennett notes,

The Multiple Drafts Model did not provide, however, a sufficiently vivid and imagination-friendly antidote to the Cartesian imagery we have all grown up with, so more recently I have proposed what I consider to be a more useful guiding metaphor: 'fame in the brain' or 'cerebral celebrity' (Dennett, 2001, p.224).

The fame in the brain metaphor makes the point that just as the moment at which a person or event became famous is not precisely datable, for some event to become famous in the brain (i.e., the transition from unconsciousness to consciousness) suffers from a similar problem when we attempt to precisely date the onset of consciousness. Consciousness, according to this view, is a competitive phenomenon where not all content can be famous at once and the only way to determine whether some event was conscious is through retrospective analysis (i.e., memory).

Dennett continues on this line of thought highlighting an exaggeration uttered by Warhol who claimed that “in the future, everybody will be famous for fifteen minutes” (Dennett, 2005, p.162). Just as a person or event couldn’t be famous for only fifteen minutes and still be considered famous (as that would be a misunderstanding of what fame is), for Dennett, some event “cannot be conscious for fifteen milliseconds” and subsequently forgotten, for that would not count as consciousness (Dennett & Akin, 2008).

Dennett is claiming that an important feature of what makes some mental event conscious, is that this event must at the very least be remembered for some length of time, and cannot be swiftly forgotten. If a competing event does not get committed to memory—at least for some time—the mental event does not count as consciousness. As Dennett writes, “our capacity to relive or rekindle contentful events is the most important feature of consciousness--indeed, as close to a defining feature as we will ever find”. Furthermore, Dennett thinks that human beings acquired habits of “self-stimulation” (i.e., reflecting upon events in the way of episodic memory) from culture, and this self-stimulation was made possible by our brains being a Joycean machine or a virtual machine, made of memes (Dennett, 2005, pgs.171-172).

Conclusion:

Chapter 3 focused on two different models of consciousness: 1) CM and 2) MDM. CM is committed to the view that there is a central location in the brain (a Cartesian theatre) where information is represented to a homunculus. Dennett’s MDM rejects that there is anything like a Cartesian theatre, instead content discriminations only need to be made once. This move is supposed to avoid the problem of an infinite regress.

As we will see in Chapter 4, CM is the basis of our folk understandings of qualia.

Although Dennett is not an eliminativist about folk psychology or intentionality (which is the subject of Chapter 3), Dennett is an eliminativist about qualia.

Dennett's Attack On Qualia:

Qualia have been thought to be an unresolvable obstacle for the scientific study of consciousness. The reason why qualia are thought to present an obstacle to science is because they are believed to be only accessible from the first person perspective. As Nagel has argued the scientific view is a third person activity that will consequently miss the first person qualitative experience since it is only revealed through introspection (Nagel, 1974; Raffman, 1995).

Dennett agrees with Nagel's first claim that the scientific perspective indeed adopts third person methodology. However, Dennett denies the second claim that consciousness is only available from the first person perspective. In Dennett's view, the method he calls heterophenomenology (phenomenology of another) allows for third person investigation that "handsomely" covers the ground that Nagel insists is missed by third person investigation (i.e., the scientific point of view).

In this Chapter I will be discussing Dennett's attack on qualia. As will be shown Dennett makes frequent use of verificationist thought experiments arguing for the elimination of qualia. The structure of this chapter will be as follows. First, I will explain Dennett's position on qualia, assessing his verificationist arguments against qualia being intrinsic, ineffable, directly introspective, and private aspects of conscious experience. Second, I will proceed to explain some of the issues at play with qualia analysing a

number of thought experiments that appear in Dennett's work. Thirdly, I will examine Dennett's attack on the Stalinesque, and Orwellian explanations of phi phenomenon and metacontrast in order to assess whether his verificationism undermines the notion of qualia being a real possibility.

Although less apparent in *Quining Qualia* (hereafter QQ), Dennett's attack on qualia is part of a broader project of attacking a view that Dennett calls Cartesian materialism, the position one adopts when one denies Descartes' dualism but keeps a special place in the brain where all inputs come together in a special structure or medium where information enter into conscious experience (Akins & Dennett, 2008). Dennett agrees that there at least seem to be qualia, what he denies is that there actually are qualia (Dennett, 1988, p.43). For Dennett, although it seems as though the experience of seeing red is an indescribable, intrinsic, and private aspect of experience, he denies that there really are any properties underlying that experience that fit the bill and therefore we ought to eliminate the idea of qualia. This view has not gone unchallenged by philosophers. For example, Flanagan agrees with Dennett that we should eliminate the qualia that Dennett has attacked, but we should not eliminate the real qualia, the way things seem to us (Flanagan, 1992, p.61).

There are four properties that Dennett ascribes to qualia which are ubiquitous in the literature, representing an underlying Cartesian materialist commitment to a unique place in the mind where everything comes together. Before I argue my point, I wish to give a brief explanation of the four properties commonly ascribed to qualia:

- Qualia are ineffable properties that cannot in principle be described no matter how well someone speaks or how open-minded the listener is.
- Qualia are intrinsic (non-relational) properties independent of other cognitive activities in the brain.
- Qualia are private properties that are beyond the ken of third person empirical science.
- Qualia are incorrigible, and directly apprehensible. That is, one cannot be mistaken about one's qualia since one has direct access to one's mental states through introspection (Dennett, 1988).

In this first section, I will begin by giving an expanded explanation of the four distinct properties that according to Dennett, philosophers ascribe to qualia, then I will consider whether Dennett's intuition pumps hold water against the four properties ascribed to qualia.

Are Qualia Ineffable?

The first property that Dennett ascribes to qualia is ineffability. By ineffable, Dennett means that qualia are supposed to be somehow indescribable. For example, people often say that colours are ineffable, since unless a subject has seen colours, it is hard to give an account of what colours look like. As Kant writes, "the person who is born blind, we can't make him see any colours" (Kant, 2006, p.61). Kant insists that unless there is a sensuous component to the reception of information about colour, then there can in fact be no experience of colour whatsoever (Kant, 2006, p.61).

More recently there have been other thought experiments that have become staples in philosophical discourse. Mostly aimed at physicalism, or functionalism, these attacks often try to draw out the intuition that a scientific explanation of consciousness leaves something out, namely qualia. One of the most famous attacks on the view that science can explain the mind without leaving something out is Frank Jackson's famous Mary's room thought experiment. Although his point was against physicalism, and in favour of epiphenomenalism (the view that consciousness is an effect and is real but has no effect on thought or action), his thought experiment also works equally well to make the point that the description of the visual spectrum that physicists give us, and its interaction with the visual system of the subject seems different from what is represented to us from the subjective, first person perspective. As Frank Jackson writes,

Mary is a brilliant scientist who is, for whatever reason, forced to investigate the world from a black and white room via a black and white television monitor. She specialises in the neurophysiology of vision and acquires, let us suppose, all the physical information there is to obtain about what goes on when we see ripe tomatoes, or the sky, and use terms like 'red', 'blue', and so on. She discovers, for example, just which wave-length combinations from the sky stimulate the retina, and exactly how this produces via the central nervous system the contraction of the vocal chords and expulsion of air from the lungs that results in the uttering of the sentence 'The sky is blue'...

What will happen when Mary is released from her black and white room or is given a colour television monitor? Will she learn anything or not? It seems just obvious that she will learn something about the world and our visual experience of it. But then it is inescapable that her previous knowledge was incomplete. But she had all the physical information. Ergo there is more to have than that, and Physicalism is false (Jackson, 1982, p.130).

The *Mary the Colour Scientist* thought experiment has been an object of contention since its introduction on the philosophical scene in 1982, and has since spawned thousands of responses, clarifications, and even an anthology titled, *There's Something about Mary* (Ludlow, et al. 2004). Jackson's thought experiment has divided philosophers into several camps. On the one hand, there are those who think that Jackson is right, and that Mary gains new knowledge that allows Mary to experience what it's like to see red, and there

are those on the other side who think that there is something wrong with the thought experiment in the first place (Dennett, 1991; Churchland, 1998, p.143).

As Dennett falls in the latter camp, it makes some sense to explain his position on the Mary's room thought experiment. For Dennett, the major difficulty with the thought experiment is how we tend to imagine the scenario, in fact, Dennett thinks that we cannot imagine what we are asked to imagine. The premise Dennett is going after is the view that we can imagine what it's like to know all of the physical information that there is about colour. We cannot in principle imagine what it's like to have all of the physical information. As Dennett protests,

The image is wrong; if that is the way you imagine the case you are simply not following directions! The reason no one follows directions is because what they ask you to imagine is so preposterously immense, you can't even try. The crucial premise is that "She has all the physical information." That is not readily imaginable, so no one bothers" (Dennett, 1991, p.399).

Note that Dennett does not think that it is merely a matter of intellectual laziness that philosophers have not imagined the case properly, it is that the thought experiment plays on a heuristic where we take a natural, but fallacious shortcut without attempting to imagine what is asked because the information is not "readily" available. Essentially people skip the step where they are asked to imagine a scenario which they cannot possibly imagine (no one knows what it's like to have all of the physical information about colour), and they leap to the intuitive conclusion that Mary would learn something new. Instead, Dennett insists that all that people actually imagine is a drop in the bucket of knowledge about what we currently know about neurophysiology of colour. Thus, people have substituted imagining something impossible for something more plausible, but not what they were supposed to imagine in the first place.

Dennett further insists that we do the thought experiment the way Jackson imagines it. Dennett gives an alternative conclusion that he thinks is equally if not more plausible. In the scenario Dennett imagines, instead of the experimenter showing Mary a red apple, and her learning some new, the experimenter presents her with a banana that he has painted blue. Instead of saying “aha, that’s what yellow looks like”, she says “aha, nice try. Bananas are yellow, not blue”. How would she accomplish such a feat without acquaintance with colour? Dennett has a reply to this question. Although we imagine the scenario as her knowing a lot, she actually would know everything there is to know about colour, as such, she would be able figure out how sensory information would affect her brain in specific ways to produce colour sensations. From there she only needs to figure out “from the inside” how red, but not yellow, would affect her visual system, how blue, but not yellow would do so as well, and eventually she would figure out how the whole visual spectrum would be represented in her colour organizational system (Dennett, 1991, pgs. 400-401). So are colour qualia as ineffable as Jackson’s thought experiment makes them out to be? Not necessarily, as Dennett has argued.

Is Dennett’s argument persuasive? Dennett is claiming not that Mary could know everything there is to know about colour, instead he is claiming that if we follow the instructions of the thought experiment, then we could get a radically different conclusion than the one reached by Jackson. What Dennett is trying to undermine is the notion that the thought experiment proves anything. As Dennett writes,

My point is not that my way of telling the rest of the story proves that Mary doesn’t learn anything, but the usual way of imagining the story doesn’t prove that she does. It doesn’t prove anything; it simply pumps the intuition that she does (“it seems just obvious”) by lulling you into imagining something other than what the premises require (Dennett, 1991, p.400).

Here Dennett is pointing out a general problem with intuition pumps, they occasionally sneak premises in (or in this case mask premises) with narrative. If we follow the premises laid down by Jackson, then she's supposed to know everything there is to know about the physical nature of colours. That's a lot of information. If we try imagining all there is to know about almost any subject whether it be chess or trees we swiftly find that we cannot really imagine the topic to its fullest breadth, and that's restricting ourselves to what is currently available as information about these subjects. Imagine the possibility of knowing everything that could be known about a subject such as colour and it becomes even more mind-boggling how much physical information there could be out there.

Interestingly Dennett does not want to emphasize this point, but rather wants to make an alternative claim that Mary could (for all we know) deduce the way a colour looks by mathematical theorem (Dennett, 2005, pgs.106-107). If colour could be deduced from a mathematical theorem this would mean that colour qualia are not ineffable, they are entirely describable by the tools of science, even the *a priori* tools. What Dennett's opponents would have to show in order to prove that Dennett is wrong is construct a mathematical proof that one cannot construct a mathematical proof deducing the colours that Mary would see.

Is Dennett's challenge unfair? Although proving a negative is uncommon it has happened in the past, and Dennett has even given us intuition pumps that design conditions under which we could prove a negative. For example, in the Able, Baker, Charlie intuition pump that Dennett gives in *Caveat Emptor*, he actually designs conditions under which we could prove a negative. Imagine there is a horse race happening where all of the horses suddenly turn off of the track, and go back to the

stables. Who won the race? The answer is no one won the race because there was no finish line. Since a necessary condition for races is a finish line of some sort or another, it could not be possible to have a race if there was not a finish line. When there are well defined boundaries and conditions for the existence of x, showing not x is always within reach.

So can we prove that Mary cannot deduce the colours she will see upon release? For all we know, it is possible to deduce the colours of the rainbow from a mathematical theorem. There's nothing built into the laws of physics which says we cannot do so.

Are Qualia Intrinsic?

When philosophers talk about the intrinsic nature of qualia, what they mean is that the perceptual representation is singular and atomic (cannot be broken down further). For example, when Schlick writes,

I observe two pieces of green paper and determine that they have the same colour. The proposition which asserts the sameness is verified... by the fact that at the same time I have two experiences of the same colour... [the proposition] cannot be reduced to any others (Schlick, 1959, pgs.92-93).

Schlick is describing the intrinsic nature of qualia which is supposed to be subjective part of experience that cannot be reduced to any other level of analysis. This immediately places philosophers who think that qualia are intrinsic (that qualia possess their properties independently of context), against those who think that qualia have relational properties that cannot be separated from the evolutionary, learning, memory, and social context in which they occur. As we will see Dennett thinks that the intuition driving the former view is false, that what we label 'qualia' are actually internal discriminations we make that cause us to behave in certain ways that cannot be separated from the context in which they occur.

In the intuition pump *Watching You Eat Cauliflower*, Dennett asks the reader to imagine subject A eating cauliflower in front of subject B. Whereas A appears to be really enjoying the cauliflower, B claims she is disgusted by cauliflower, and cannot imagine how A enjoys that taste. The conclusion drawn by B is that cauliflower must taste different to A to explain the difference in behaviour. Dennett elaborates on the example, by asking the reader to imagine an intrapersonal account where A drinks orange juice at time t , then proceeds to eat pancakes at time t' which makes the orange juice taste different than it did at time t . But after a cup of coffee, the orange juice goes back to tasting as it did at time t (Dennett, 1988, p.44-45).

Intuitively we might think there is a difference that can be discovered between when A tasted orange juice prior to eating pancakes that can be analysed separately from the way orange juice tastes after eating pancakes, but Dennett thinks this intuition is misleading because it attempts to isolate qualia from everything else going on within the system (Dennett, 1988, p.45). Why does Dennett believe trying isolate qualia is a mistake? Let's use colour qualia as an example to illustrate what Dennett is thinking.

Consider first the interpersonal inverted spectrum thought experiment. As the story goes, it is entirely possible that we could both have our colour qualia inverted in such a way that when I see red, you see my green and vice versa. For example, imagine we are both looking at a tulip, and I say "what beautiful red pedals", and you agree with me. Even though we both agree that the tulip is red and beautiful, we are not experiencing the same colour qualia. Since there would be no difference in our verbal behaviour that would constitute evidence that we are having distinct qualitative experiences, it was argued that there was no way of verifying whether inverted spectrums were a real

possibility and therefore, inverted spectrums were utter non-sense or “meaningless” as Schlick puts it (Dennett, 1988, p.49; Shoemaker, 1982, p.357; Schlick, 1959, p.93).

What Dennett wants to deny is that there is any possibility of discussing the way something tastes, smells, feels, apart from the organism’s evolutionary, and learning history as well as its propensity to make verbal judgments about discriminations it has made (Thompson, p.37). Instead Dennett proposes that we look at the functional processes that produced those judgments. We consider only the information bearing system that is responsible for the seeming to be qualia. Take for example Dennett’s intuition pump of a wine tasting machine. After pouring the wine into the contraption it makes a number of sensory discriminations that an experienced wine taster can make by chemically breaking down the wine into its constituent parts and producing as output in the form of a description of the wine with all of the predictive accuracy of the experienced wine-taster. If we think that this output could not be produced without any of the qualia that we enjoy, then this is the atomic, and intrinsic property that Dennett wants to deny as existing (Dennett, 1988, p.46).

Are Qualia Private? Are Qualia Incorrigible?

There are two senses in which the term ‘representation’ is used. First, there are representations in the sense that we use them today to refer to the process of the brain tracking objects in the real world (Thompson, 2009, p.30). Dennett does not deny that representations exist in this sense, in fact, the multiple drafts model presupposes it by treating the brain as a massive parallel processor that tracks sensory information about objects through real space, and real time. When Dennett denies that qualia exist, he is denying that there are real Seemings, that there are representations in the sense of

Descartes (Thompson, 2009, p.30). A real Seeming is a representation in that it represents the object of the real world to a homunculus in the mind/brain. Hence, why Dennett labels this place in the mind/brain the Cartesian theater, because the representation is played for the privileged audience, namely the Cartesian homunculus. What Dennett wants to deny is that there is any motivation for such a theater (metaphorically and literally) that imposes a dichotomy between inner and outer experience. Instead, Dennett wishes to say there is just one world that has objects causally interacting. The brain—according to Dennett’s view—is as much a part of this world as stars in the sky are a part of our universe. Thus, although there seem to be qualia that exist outside the realm of inter-observer experience, there really are no such private properties, therefore, qualia should be eliminated with other theoretically unnecessary entities such as phlogiston. Not only does Dennett want to deny that qualia are private and inaccessible from inter-observer experience, Dennett wishes to deny that qualia are incorrigible, that since we are in a privileged position as observers of our own conscious entities that we could not be mistaken about our own qualia. Since Dennett’s intuition pumps are deeply interwoven I want to look at both of these properties at the same time explaining why Dennett thinks these two properties should be eliminated.

As Dennett poses the problem,

I cannot imagine, will never know, could never know, it seems, how Bach sounded to Glenn Gould. (I can barely recover in my memory the way Bach sounded to me when I was a child.) And I cannot know, it seems what it is like to be a bat (Nagel, 1974), or whether you see what I see, colour wise, when we look up at a clear ‘blue’ sky (Dennett, 1994, p.45).

As mentioned in the section on ineffability, there is a distinction throughout Dennett’s work on qualia, between the way things seem, and the way they actually are. Kant makes much of this distinction in his monumental work, the *Critique of Pure Reason*, where he

discusses the difference between the way things appear to us, and the way things are in-themselves. Whereas Kant argued for transcendental idealism, the view that we are epistemically cut-off from how things are in-themselves by the appearance of things, and therefore, we can never know the true nature of say the colour of a rose, we can only know its mode of presentation or the way it appears; Dennett endorses a form of epistemic verificationism, where the way things appear to us is authoritative of only appearances or as Dennett, and Brook put it—Seemings—but we are not authorities of how things actually are (Brook, 2002, p.41-61). Science is the authority of how things are, and although science could be wrong about our own first person experience, it is far more likely to be correct than first person phenomenological inventory where we introspect for how our consciousness actually is.

This is not to say Dennett is offering a normative position for how we should blindly accept the dictates of science, rather, Dennett is contending that we should take the science seriously, as well as the first person experience as two distinct levels of analysis that can be corrected by further facts of the matter. One of Dennett's favourite examples is the story of Chase, and Sanborn who are two Maxwell House taste testers who both agree that they no longer like the coffee they have been taste testing, but give two seemingly different accounts of how the mechanisms behind their first person experiences of their "qualia" actually are. As Chase says in the intuition pump,

I hate to admit it, but I'm not enjoying the work anymore. When I came to Maxwell House six years ago, I thought Maxwell House coffee was the best tasting coffee in the world. I was proud to have a share in the responsibility for preserving that flavour over the years. And we've done our job well; the coffee tastes just the same today as it tasted when I arrived. But... I no longer like it! My tastes have changed. I've become a more sophisticated coffee drinker. I no longer like that taste at all (Dennett, 1994, p.52).

Then Sanborn responds with the following dialogue,

When I arrived here, shortly before you did, I, like you, thought Maxwell House coffee was the tops in flavour. And now I, like you, really don't care for the coffee we're making. But my tastes haven't changed, my... tasters have changed. That is, I think something has gone wrong with my taste buds or some other part of my taste-analyzing perceptual machinery. Maxwell House coffee doesn't taste to me the way it used to taste if only it did, I'd still love it, for I still think that taste is the best taste in coffee. Now I'm not saying that we haven't done our job well. You other tasters all agree that the taste is the same, and I must admit that on a day-to-day basis I can detect no change either. So it must be my problem alone. I guess I'm no longer cut out for this work (Dennett, 1994, p.52).

The upshot of the thought experiment is that although we can often give very different accounts of what is actually the case from the appearances of things, these appearances alone do not offer the evidence that we need in order to confirm or deny the two competing conclusions that Chase, and Sanborn come to. We can agree with Chase that it seems to him that his judgements have changed, we can also agree that it seems to Sanborn as if his perceptions have changed, but in order to confirm or deny either of these conclusions, we need to look to a third person account to be the arbiter between these two competing explanations. If third person analysis cannot perform an experiment in principle to decide between the two competing explanations, then there is no difference between the two explanations, it's a distinction without a difference.

In Dennett's view, if a question cannot be answered in principle, then there is a problem with the question in the first place. The problem in the Chase, and Sanborn intuition pump is the evidential dependency on the utterances of the subjects, and the lack of third party neurological data that could in principle falsify the theories. Since the empirical data could be used to support either theory, the evidence is thus largely underdetermined. In other words, Dennett is making two claims: first, that all of the data cannot simply confirm both hypotheses, the data must also be capable of proving that the hypotheses are false if indeed they are false, and second, that there must be a measurable difference between the two explanations, if not, then the explanations are meaningless.

The former of the two claims avoids the confirmation excesses that lead to unjustified inductive conclusions by ruling out the usage of data that would only support a hypothesis. The latter is a verificationist argument to the in the sense that it's both an anti-realist move, and derives the meaningfulness of the explanation from observation (or in this case, subtracts meaning when two competing explanations cannot be decided by any observation whatsoever).

Not everyone will be persuaded by this argument against there being any fact whatsoever about the Chase and Sanborn case, especially those who are realists about phenomenon beyond observation, and those who see neuroscience as capable of mapping the functions associated shifting aesthetic judgments, and the mechanisms that underlie perceptual shifts.

Phi Phenomenon and Verificationism:

In this section, I will discuss how Dennett's view on phi phenomenon is another example of the distinction without a difference move and how it directly parallels the argument presented by Dennett in the Chase and Sanborn intuition pump. Furthermore, I will explain how Dennett's argument against CM explanations of phi phenomenon poses an issue for those who think that qualia are incorrigible.

In order to accomplish this task, I must first explain what phi phenomenon is. Phi phenomenon is a perceptual illusion caused by the rapid succession of stationary stimuli such as pictures (or in Dennett's example, light) in which the onset of the stimuli causes the subject to perceive an apparent motion between the two stationary stimuli. There are many instances in which apparent motion occurs (movies are an example), however, one example Dennett uses throughout CE is from an experiment performed by Kolars and

Grünau is known as phi phenomenon (Dennett, 1991, p.114). In the experiment, two lights (a green one and a red one) are flashed in a non-synchronized pattern, where the red light flashes first and then the green light flashes second in rapid succession, the subject observing the pattern often judges that the red light moved and then abruptly morphed into the green light. One question that was raised by Nelson Goodman was “how are we able to fill in the spot at the intervening place-times along a path running from the first to the second flash before the flash even occurs”? (Dennett, 1991, p.114)

Dennett presents two alternative explanations for phi phenomenon that he calls the Orwellian and Stalinesque interpretations, but as Dennett attempts to show neither works as a good explanation for phi phenomenon because all of the data we could acquire would confirm both explanations. Since no information from the first person and third person perspectives could distinguish either explanation from one another, Cartesian materialism is taken to be false (Bloomfield, p.211; Schneider, p.315).

According to the Orwellian interpretation the subject perceives the stimulus correctly, but then constructs a false memory of the event. Since the false memory overwrites the original perception of the stimulus (and does so quickly) the subject’s reports are always based on the false memory (Dennett, 1991, pgs.120-121). According to the Stalinesque interpretation, the brain has a time-delay for conscious experience where the onset of consciousness is delayed long enough for a perception editor to fill in the facts that were not in the stimulus itself. Thus, the subject’s memory is correct, but her consciousness is mistaken (Dennett, 1991, p.120).

Dennett wants to deny that neuroscience can give us fine-grained enough map of how function superimposes onto structure. If Dennett is correct, just as he thought there

was no way *in principle* to distinguish Chase and Sanborn explanations about how coffee tasted then and now, Dennett thinks there cannot be any way *in principle* to distinguish an Orwellian from a Stalinesque explanation. Since he concludes that there simply are not any facts that can be attributed to Stalinesque/Orwellian and Chase/Sanborn explanations Dennett is adopting an anti-realist position that places him squarely within the verificationist camp.

Another parallel between the Chase/Sanborn and Stalinesque/Orwellian distinctions is that there is no way for the subject to verify with his/her first person experience whether one explanation or the other is correct. Just as one cannot introspect to discover whether one's perceptual machinery has changed or one's aesthetic judgments have changed, one cannot introspect to discover why the red dot seems to morph turning green before the light possibly could have turned on. Since it's impossible to introspect on how one's neurons are being affected by stimuli, a subject must submit herself to third person verification which as Dennett points out, immediately eliminates the possibility of incorrigibility of qualitative states.

But is Dennett's verificationist attack on the Stalinesque and Orwellian explanations of phi phenomenon compelling? Recall that for Dennett there are two stories being told that although upon appearances they seem to give a different story of what's going on inside the subjects of the phi phenomenon, the explanations are only different in respect to where they place the divide between consciousness and unconsciousness, which is a verbal difference and no more.

Metacontrast:

In CE Dennett discusses the metacontrast experiments that he claims have snuck in Cartesian materialist explanations in the explanations of the results. One metacontrast experiment that Dennett discusses has two visual stimuli flashed in rapid succession before the observer. The first stimulus is a small circle that fits inside a ring with a hole in it. The second stimulus is the ring with the hole. When the two stimuli are flashed very quickly in front of the observer, the subjects—when probed—claim that they saw a ring, but they do not claim to see a small circle. Only when they are asked to guess if there were one or two stimuli, do the subjects (at better than chance) confirm that there were two stimuli (Dennett, 1991, p.141).

Like phi phenomenon and the Chase and Sanborn thought experiment, the metacontrast experiments have two competing explanations for the behaviour of subjects. On the one hand, the Stalinesque interpretation of metacontrast effect claims that the subject is prevented from ever becoming conscious by the masking stimulus. Since the ring prevents the circle from entering the privileged circle of consciousness, the experiment is interpreted as showing that stimuli can have effects on the behaviour of the organism without her ever becoming conscious of the stimuli. On the other hand, the Orwellian interpretation of metacontrast effect takes the subject to have been conscious of the first circle but to have her memory swiftly altered by the second stimulus so as to nearly obliterate the memory of the first stimulus (Dennett, 1991, p.142).

Dennett thinks that both explanations can account for all of the current data, they can also account for all of the data we could foreseeably retrieve in the future. Since both sides could not be able to agree on a crucial experiment that would be able to falsify one

or both of the theories, there is no way to settle the dispute (Dennett, 1991, p.142). Again, the problem for Dennett is there is no way *in principle* of constructing an experiment that could settle the issue if we stick to these two explanations. Hence, the metacontrast effect requires an alternative explanation that does not depend on unresolvable explanations, to Dennett we instead need an explanation that can be empirically settled.

Analysis:

In this section I will give some analysis of the conclusions Dennett comes to from the intuition pumps and experiments he gives. First, let us consider the case of Chase and Sanborn, is there really no measurable difference between either explanation?

For example, what if we measured the taste receptors of Sanborn over time and discovered that indeed his taste receptors were slowly degenerating? In that case, we would have third person evidence corroborating his view that Maxwell House coffee doesn't taste the same as it used to. Moreover, if we found no difference in his taste buds or taste testing machinery we might suspect that he was mistaken. Of course this would mean that Sanborn's qualia are not private, since the only way to discern whether his conclusion about his taste buds being altered was to submit Sanborn to third person testing and analysis.

It is also possible that Sanborn could perform experiments on his own discriminative capacities and infer that his taste buds have degenerated by testing his lowest thresholds for tasting various substances. But he would have at least needed to test himself before the degeneration happened as a benchmark for later testing.

If this procedure was possible (and it is) this would mean that there is a fact of the matter about whether Sanborn's taste buds have altered over time. Thus, at least Sanborn's case is verifiable.

In the case of Chase, if we operationalize coffee drinking sophistication to the measurable action of drinking anything but Maxwell House (and incorporating variety into our measurements) over time and notice that he has been drinking less Maxwell house outside of work and has been drinking imported fair trade coffees from Columbia, then we could have a measurable difference that would suggest he had become a more sophisticated coffee drinker.

If the Chase and Sanborn case presents genuinely verifiable differences, that makes Dennett wrong that there is no way in principle to decide between the two cases. But does this pose a problem for Dennett's argument? Not at all. It just shows that qualia are neither private, nor ineffable, nor indescribable from third person science which suggests elimination of these properties of qualia.

What about phi phenomenon? At ever increasingly small time scales, it is challenging to measure whether memory has been laid down or not. Since we cannot at this point measure the precise moment in which memory has been laid down, there would be no way in practice to discover whether Stalinesque or Orwellian memory alteration happened. Furthermore, since both explanations agree that there is memory revision, they just disagree about where it is placed in time, and measurement itself takes time, at such a small time interval it does not seem likely that there could be a technology that could make such a fine distinction. The same issues would also be at play in the metacontrast

example since they operate at similar time scales. In the end, I do cautiously agree with Dennett's assessment of phi phenomenon.

That being said, Dennett and I for that matter could be wrong. What has seemed impossible in the past, has turned out to be possible with advancement in technologies. If we asked a person who lived before we had an understanding of the speed of light, radio waves, etc., whether it would be possible to send a message almost instantaneously across the Atlantic from Canada to England, the person would think it impossible in practice if not in principle. Nothing in their experience would give them any indication that there was a way in which to accomplish this task obeying the laws of nature. Thus, although I agree with Dennett's conclusions about phi phenomenon and metacontrast being impossible in principle to distinguish between Stalinesque and Orwellian explanations for phi phenomenon and metacontrast experiments, I think it possible that Dennett and I could be shown to be wrong.

Conclusion:

Dennett's intuition pumps and experiments make a compelling case for eliminating the concept of qualia (and I share his sentiments), qualia are too messy to work with and I think I know why. There does not seem to be any consensus on what qualia are except that they all refer to sensory experiences or how things feel (e.g., the taste of an orange for example or the smell of a rose). But because qualia captures so many different qualitative experiences in one term, the concept becomes far too general to establish consensus in regard to the properties they could have. Some examples of "qualia" are going to be more amenable to verification (such as the Chase and Sanborn case) than others. Moreover, there will be "qualia" that won't be verifiable as many of

Dennett's thought experiments suggest. Thus, we end up with inevitable contradictions whether we accept or reject Dennett's conclusions because the concept itself is tangled up. Ultimately I agree with Dennett's eliminativism towards qualia for the reasons that qualia are far too messy, and are conceptually too general to be of any utility to scientific or philosophical investigation, since qualia picks out far too many different kinds of phenomena that are more or less amenable to scientific and philosophical investigation.

Conclusion of Thesis:

In Chapter 1, I wrote a brief but succinct history of verificationism that was intended to describe the problems and faces of those who pursued a verificationist criterion of meaning. Some of the figures discussed were Kant and Hume from the enlightenment period, both of whom effected the course that the logical positivists would later take with their verificationist criterion of meaning. I discussed the distinction between verification in practice and verification in principle as well as some of the difficulties verificationism faced which eventually led to its downfall.

Chapter 2, was dedicated to introducing Dennett's thoughts on the intentional stance and developing some of the familiar criticisms of the intentional stance as well as some of my own. Since the intentional stance finds itself in the realm of folk psychology because people do in fact use the intentional stance to make predictions and explain other people's behaviour, there's a further question as to whether folk psychology is a theory—as Paul Churchland suggests while Dennett dissents—which is explored in the chapter. I also question whether Dennett's realism with regard to intentional states, specifically beliefs, is strong enough.

By exploring Dennett's verificationist views on the intentional stance and his instrumentalism in regard to intentional states I found myself exploring Dennett's attack on Cartesian materialism and his subsequent replacement with the model of consciousness called the multiple drafts model later renamed fame in the brain.

In chapter 4, I explored Dennett's crucial verificationist attack against Cartesian materialism, his attack on qualia. Here I found myself in agreement with Dennett's eliminative materialism against qualia for a number of reasons. The first reason I am in agreement with Dennett is precisely the reason he gives in his introduction to QQ, that the notion of qualia is much confused and there appears to be very little possibility of dissolving this confusion. By and large my contribution in this chapter has been to collect the various verificationist arguments that Dennett makes throughout his attack on Cartesian materialism and exposit how they work to develop a clearer picture of Dennett's verificationist moves.

Potential Future Research on the Topic:

There are a number of questions that could be addressed that I think would yield some fruitful discussion. More thought experiments on qualia and especially intentionality could be examined and dissected.

Discussions about Dennett's instrumentalism could be made more clear by building a more comprehensive list of Dennett's critics who read Dennett to be a sort of instrumentalist. Further, a more comprehensive list of those who give non-instrumentalist readings of Dennett would improve the project by producing better discussion. More detailed descriptions of different styles of instrumentalism and what sort of

instrumentalist Dennett could be described as would also yield fruitful analysis. A more detailed analysis of different flavours of realism would be fruitful in my discussion of Dennett's intentional strategy and also a discussion of different styles of anti-realism would improve the project's depth of analysis.

In Chapter 1, I would add more of a discussion on Wittgenstein and his contributions to the verificationist project (which were perhaps unintentional), as well as the tension of his philosophy with the overall project of verificationism. I would also add more recent authors who have defended verificationist views such as Michael Dummett and Bas Van Fraassen.

References:

Akins, K. (Ed.). (1996). *Perception* (Vol. 5, Vancouver Studies in Cognitive Science). New York, NY.: Oxford University Press.

Akins, K. (1996). *Lost the plot? Reconstructing Dennett's multiple drafts theory of consciousness*. *Mind & Language*, 1(2), 1-43.

Ayer, A. J. (1952). *Language truth and logic*. New York, NY.: Dover.

Ayer, A. J. (1959). *Logical positivism*. New York, NY.: Free Press.

Ariew, R. (2014). Pierre Duhem. *In Stanford Encyclopedia of Philosophy*.

Brook, A., & Ross, D., & Thompson, D. (2000). *Dennett's philosophy a comprehensive assessment*.: MIT Press.

Brook, A., & Stainton, R. (2000). *Knowledge and mind a philosophical introduction*. Cambridge, Mass.: MIT Press.

Brook, A., & Ross, D. (2002). *Daniel Dennett*. Cambridge: Cambridge University Press.

Brook, A., & Ross, D. (2002.). *The appearance of things*. In A. Brook (Ed.), Daniel Dennett. Cambridge University Press.

Bechtel, W., & Graham, G. (Eds.). (1999). *A companion to cognitive science*. Malden, Mass.: Blackwell Publishing.

Berlin, I. (1977). *Hume and the Source of Anti-Rationalism*. Retrieved from http://berlin.wolf.ox.ac.uk/published_works/ac/hume.pdf

Berlin, I., & Gray, J. (1999). *The roots of romanticism* (H. Hardy, Ed.). Princeton, NJ.: Princeton University Press.

Block, N. (2007). *Consciousness, function, and representation*. Cambridge, Mass: MIT Press.

Bloomfield, P. (1998). *Dennett's misrememberings*. *Philosophia*, 26(1).

Brown, J. (1991). *The laboratory of the mind*. Routledge.

Bunnin, N. (2003). *The Blackwell companion to philosophy* (2nd ed.). Oxford: Blackwell.

Cain, M. (2002). *Fodor: Language, mind, and philosophy*. Cambridge, UK: Polity.

- Creath, R. (2014). Logical Empiricism. In *Stanford Encyclopedia of Philosophy*.
- Chalmers, D. (1996). *The conscious mind: In search of a fundamental theory*. New York: Oxford University Press.
- Churchland, P. M., & Churchland, P. S. (1998). *On the contrary: Critical essays, 1987-1997*. Cambridge, MA.: MIT Press.
- Dahlbom, B. (1997). *Dennett and his critics: Demystifying mind.*: Oxford, UK: Blackwell.
- Dawson, M. R. (n.d.). *Mind, body, world: Foundations of cognitive science*. Alberta: AU Press.
- Dennett, D. C. (1969). *Content and consciousness*. London: Routledge & K. Paul.
- Dennett, D. C. (1978). *Current issues in the philosophy of mind*. *American philosophical quarterly*, 15(4), 249-261.
- Dennett, D. C. (1982). *Comments on Rorty*. *Synthese*, 53(2), 349-356.
- Dennett, D. C. (1981). *Brainstorms.*: MIT Press.
- Dennett, D., & Hofstadter, D. (Eds.). (1982). *The minds I: Fantasies and reflections on self and soul*. New York, NY.: Bantam Books.
- Dennett, D. C. (1987). *The intentional stance*. Cambridge, Mass.: MIT Press.
- Dennett, D. (1988). *Quining qualia*. In A. Marcel (Ed.), *Consciousness in Contemporary Society* (pp. 42-78). Oxford.
- Dennett, D. C. (1991). *Consciousness explained*. Boston.: Little Brown Book.
- Dennett, D. C., & Kinsbourne, M. (1992). *Time and the observer*. *Behavioural and Brain Sciences*, 15(2), 183-247.
- Dennett, D. C. (1993). *Caveat emptor*. *Consciousness and Cognition*, 2(1), 48-57.
- Dennett, D. C. (1993b). *The message is there is no medium*. *Philosophy and Phenomenological Research*, 53(4), 919-931.
- Dennett, D. C. (1996). *Seeing is believing: Or is it?* (K. Akins, Ed.). In *Perception* (Vol. 5, pp. 158-172). New York, NY: Vancouver Studies in Cognitive Science.
- Dennett, D. C. (1998). *Brainchildren essays on designing minds*. Cambridge, Mass.: MIT Press.
- Dennett, D. C. (2001). *Are we explaining consciousness yet?*. *Cognition*, 79, 221-237.
- Dennett, D. (2005). *Sweet dreams philosophical obstacles to a science of consciousness*. Cambridge, Mass.: MIT Press.

- Dennett, D. C & Akins, K. (2008). The multiple drafts model. *In Scholarpedia*.
- Dennett, D. C. (2013). *Intuition pumps and other tools for thinking*. W.W. Norton.
- Dennett, D. C. (2013). *Real consciousness*. In A. Revonsuo & M. Kamppinen (Eds.), *Consciousness in Philosophy and Cognitive Neuroscience*. Psychology Press.
- Flanagan, O. (1992). *Consciousness reconsidered*. Cambridge, Mass.: MIT Press.
- Flanagan, O. (1991). *The science of the mind*. Cambridge, Mass.: MIT Press.
- Frankish, K., & Ramsey, W. (2012). *Cambridge Handbook of Cognitive Science*. Cambridge: Cambridge University Press.
- Friedman, M. (1999). *Reconsidering Logical Positivism*.: Cambridge University Press.
- Graham, G. (2015). Behaviourism. In E. Zalta (Ed.), *Stanford Encyclopedia of Philosophy*.
- Greenwood, J. D. (1991). *The future of folk psychology: Intentionality and cognitive science*.: Cambridge University Press.
- Hanna, R. (2014). Kant's Theory of Judgment. In *Stanford Encyclopedia of Philosophy*.
- Hardcastle, V. (1995). *Locating consciousness*. Amsterdam: J. Benjamins Pub.
- Hardin, C. L. (1988). *Color for philosophers: Unweaving the rainbow*.: Hackett.
- Hempel, C. G. (2000). *Carl G. Hempel: Selected philosophical essays*.: Cambridge University Press.
- Howard, R. (2016). Dualism. In *Stanford Encyclopedia of Philosophy*.
- Huemer, M., & Zalta, E. (2011). Sense-Data. In *Stanford Encyclopedia of Philosophy*.
- Hume, D. (1985). *A Treatise Concerning Human Nature*. England: Penguin Classics.
- Hume, D. (1748). *An inquiry concerning human understanding*.: The Library of Liberal Arts.
- Jackson, F. (1982). *Epiphenomenal Qualia*. *Philosophical Quarterly*, 32(127), 127-136.
- Joergensen, J. (1951). The development of logical empiricism. In *Foundations of the Unity of Science Toward an International Encyclopedia of Unified Science*, 2(9). University of Chicago Press.
- Kant, I. (1781-1787). *Critique of pure reason* (N. K. Smith, Ed.).: Macmillan.
- Kant, I., Ed. Louden, R. (2006). *Anthropology from a pragmatic point of view*. New York, New York: Cambridge University Press.
- Kitcher, P. (1990). *Review*. *Philosophical Review*, 99(1), 126-128.

- Ludlow, P., Yuljin, N., & Stoljar, D. (2004). *There's something about Mary: Essays on phenomenal consciousness and Frank Jackson's knowledge argument*. Cambridge, Mass.: MIT Press.
- Lycan, W. (2003). *Philosophy of Mind* (E. James, Ed.). In N. Bunnin (Ed.), *The blackwell companion to philosophy* (2nd ed., pp. 191-192). UK: Blackwell.
- Macdonald, G. (2010). A. J. Ayer. In E. Zalta (Ed.), *In Stanford Encyclopedia of Philosophy*.
- Marcel, A. J. (1988). In *Consciousness in contemporary science*. Oxford: Oxford University Press.
- Metzinger, T. (1995). *Conscious experience*. Paderborn: Schöningh/Imprint Academic.
- Misak, C. (1995). *Verificationism: Its history and prospects.*: Routledge.
- Murat, A. (2015). The Language of Thought Hypothesis. In E. Zalta (Ed.), *In Stanford Encyclopedia of Philosophy*.
- Nagel, T. (1974). *What Is It Like to Be a Bat?* *The Philosophical Review*, 83(4), 435-450.
- Noe, A., & Thompson, E. (Eds.). (2002). *Vision and mind: Selected readings in the philosophy of perception*. Cambridge, Mass.: MIT Press
- Okasha, S. (2002). *Philosophy of science: A very short introduction*. Oxford: Oxford University Press.
- Oberdan, T. (2013). Moritz Schlick. In *Stanford Encyclopedia of Philosophy*.
- Platchias, D. (2011). *Phenomenal consciousness: Understanding the relation between neural processes and experience*. Stocksfield: McGill Queen Press.
- Popper, K. (1959). *The logic of scientific discovery.*: Routledge Classics.
- Quine, W. V., & Ullian, J. S. (1978). *The web of belief* (2nd ed.). New York, NY: Random House.
- Raffman, D. (n.d.). *On the Persistence of Phenomenology*. In T. Metzinger (Ed.), *Conscious experience* (pp. 297-307). Imprint Academic.
- Ramsey, W. (2013). Eliminative Materialism. In *Stanford Encyclopedia of Philosophy*.
- Ravenscroft, I. (2010). Folk Psychology as a Theory. In E. Zalta (Ed.), *In Stanford Encyclopedia of Philosophy*.
- Rey, G. (1997). *Contemporary philosophy of mind: A contentiously classical approach*. Cambridge, Mass.: Blackwell.

- Rorty, R. (1998). *Truth and progress: Volume III*. Cambridge: Cambridge University Press.
- Santrock, J. W., & Mitterer, J. O. (2005). *Psychology 2*. Toronto: McGraw-Hill Ryerson.
- Schlick, M. (n.d.). Positivism and Realism. In A. J. Ayer (Ed.), *Logical positivism* (pp. 82-107). New York, NY: Free Press.
- Shoemaker, S. (1982). *The Inverted Spectrum*. *The Journal of Philosophy*, 79(7), 357-381.
- Shoemaker, S. (1990). *Reviewed work: The intentional stance*. *The Journal of Philosophy*, 87(4), 212-216.
- Searle, J. (2003). *Contemporary philosophy in the United States*. In *The Blackwell Companion to Philosophy* (2nd ed.). Oxford: Blackwell.
- Strawson, P.F., 1966, *The Bounds of Sense: An Essay on Kant's Critique of Pure Reason*, London: Methuen.
- Thompson, D. (2009). *Daniel Dennett*. London: Continuum.
- Tye, M., & Zalta, E. (2015). Qualia. In *Stanford Encyclopedia of Philosophy*.
- Uebel, T. (2014). Vienna Circle. In E. Zalta (Ed.), In *Stanford Encyclopedia of Philosophy*.
- Wright, E. (2008). *The case for qualia*. Cambridge, Mass: MIT Press.
- Zawidzki, T. (2007). *Dennett. One World*.

