

5-2009

"Turning the Fly Around" The Relationship between Wittgenstein's Discussion of Meaning and the Self: An Exegesis and Defense

Sean W. Dalby
College of William and Mary

Follow this and additional works at: <https://scholarworks.wm.edu/honorsthesis>



Part of the [Philosophy Commons](#)

Recommended Citation

Dalby, Sean W., "Turning the Fly Around" The Relationship between Wittgenstein's Discussion of Meaning and the Self: An Exegesis and Defense" (2009). *Undergraduate Honors Theses*. Paper 319.
<https://scholarworks.wm.edu/honorsthesis/319>

This Honors Thesis is brought to you for free and open access by the Theses, Dissertations, & Master Projects at W&M ScholarWorks. It has been accepted for inclusion in Undergraduate Honors Theses by an authorized administrator of W&M ScholarWorks. For more information, please contact scholarworks@wm.edu.

Turning the Fly Around

**The Relationship between Wittgenstein's Discussion of Meaning and the Self: An
Exegesis and Defense**

A thesis submitted in partial fulfillment of the requirement
for the degree of Bachelors of Arts in **Philosophy** from
The College of William and Mary

by

Sean Wesley Dalby

Accepted for _____
(Honors, High Honors, Highest Honors)

Professor Matthew Haug, Director

Professor James Harris

Professor Rob Leventhal

Williamsburg, VA
April 29, 2009

Contents

Acknowledgement	3
Abbreviations	4
Introduction:	5
Assumptions and Justifications	6
Outline	8
Chapter One:	
The Rule-Following Problem	11
I. The Initial Problem	14
II. The Dispositional Response	16
III. Frege's and Russell's Responses	20
Chapter Two:	
Uprooting Old Assumptions and Planting New Meaning	29
I. A Revolutionary Rejection	30
II. Primitive Language Games and Thought Experiments	35
III. Extending Our Results	42
Chapter Three:	
The Private Language Argument and the Problem of Other Minds	55
I. The "Private" Language Argument	55
II. Descartes' Error and the New Status of "Mental States"	64
Chapter Four:	
The Dissolution of Solipsism	78
I. The Traditional Assumption about the Self :Its Advocates	79

II. “But What do you Mean?”	82
Epilogue: Final Thoughts and Suggestions for Future Research	89
Bibliography	95

Acknowledgement

I would like to take this opportunity to thank Matt Haug for his instrumental help and guidance throughout my research for this essay. Conversations with him grounded my thoughts and focused my goals in ways that would have been impossible for me to do alone. Every good thought and argument below came from my discussions with him, and I thank him for having the patience and insight in our conversations needed to make this paper possible.

Abbreviations

PI = Philosophical Investigations

BB = The Blue and Brown Books

RFM = Remarks on the Foundations of Mathematics

Z = Zettel

Introduction

Language is a labyrinth of paths. You approach from *one* side and know your way about; you approach the same place from another side and no longer know your way about. (PI, 203)

Our language can be seen as an ancient city: a maze of little streets and squares, of old and new houses, and of houses with additions from various periods; and this surrounded by a multitude of new boroughs with straight regular streets and uniform houses. (PI, 18)

While language may be ancient and labyrinthine, Wittgenstein's metaphors retain their vision when applied to philosophy in general. Together they serve as an enlightening depiction of historical approaches to philosophical problems. Many philosophical problems have been abandoned and fallen into decay; others apparently solved. Some have taken new forms and others have been only recently discovered. Careful study of seemingly unrelated fields of philosophy reveals a deep interconnection between them all, which gives rise to much commerce and growth in the discipline as a whole. While we are free to choose our entrance to this bustling metropolis of philosophical problems, we must tread carefully to avoid a hasty choice of route into town, lest we find ourselves lost in a bad neighborhood. Hopefully our decision to enter via the philosophy of language will not lead us astray.

There are two important preliminary questions that deserve attention before any trip: where do we plan on going and why are we taking this route as opposed to others. Accordingly, the first purpose of this introduction will be to give a general justification of adopting the route of the philosophy of language. Its second purpose will be to describe the structure of the arguments below. This section gives a clear statement of the overall project of this essay and lays out the relationship between the chapters.

Assumptions and Justifications

How our words possess meaning is one of the premiere problems within the philosophy of language, and seems to be of fundamental importance for philosophy as a whole. Language is the medium through which all philosophical problems are expressed, and if the articulation of any particular problem can be shown to be incoherent or meaningless, we can re-label the problem as a pseudo-problem and dismiss it accordingly. Throughout history, many philosophers have adopted varying versions of this method, but, as the title suggests, the philosophical grandfather of this essay is Wittgenstein. Indeed, while he distances himself in his later writings from logical positivism as well as his earlier thought, he embraces this general methodology explicitly:

It is the business of philosophy, not to resolve a contradiction by means of a mathematical or logico-mathematical discovery, but to make it possible for us to get a clear view of the state of mathematics that troubles us: the state of affairs *before* the contradiction is resolved. (And this does not mean that one is sidestepping a difficulty.)

The fundamental fact here is that we lay down rules, a technique for a game, and that then when we follow the rules, things do not turn out as we had assumed. That we are therefore as it were entangled in our own rules.

This entanglement in our rules is what we want to understand (i.e. get a clear view of).

It throws light on our concept of *meaning* something. (PI, 125)

One of the primary functions of philosophy consists in explaining how our words possess meaning; doing so provides a clear view of what exactly it is that troubles us. That this is the *sole* function of philosophy is not a position of Wittgenstein's that I wish to defend. It seems to require a robust philosophical argument to support itself and so, needless to say, I find the former weaker claim much more immediately plausible. In any case, explaining how our words possess meaning often has the power of showing us what the true nature of our problems is. More importantly, the upshot of this assumption entails that we might discover something in this process that reveals our previous approaches to our problems

as fallacious. For example, once we discover that we can not really *mean* that mental states are “objects” in the mind of a person, we quickly see how traditional Cartesian assumptions about the mind become problematic. However, we can start looking for what they might in fact be in light of our philosophical reflections on language. The discussion below relies broadly on this assumption. Using this as a starting point, it is my intention to connect a specific novel conception of meaning to inform our understanding of mental states and the self, which, if done properly, will give credence to this method. More generally, this method assumes that the insights of philosophy of language can inform our discussions in metaphysics and the philosophy of mind specifically.

Outline

The central claim of this essay argues that there is an important relationship between Wittgenstein's discussion of rule following and meaning with his discussion of solipsism. Historically, Wittgenstein always considered solipsism an important target. Gordon Baker and Peter Hacker even bring us evidence that the fly in Wittgenstein's infamous fly-bottle *is* the solipsist:

‘The solipsist flutters and flutters in the flyglass, strikes against the walls, flutters further. How can he be brought to rest?’ Note that in MS 165 Wittgenstein states explicitly that the discussion of a private language *concerns the problems of idealism and solipsism*. (Baker and Hacker 1984, p. 23)

Extensive literature exists concerning Wittgenstein's discussion of rule-following (e.g. Kripke 1982, Boghossian 1989, 1994, Baker and Hacker 1984, etc.) and much has been written on his discussion of solipsism later in the *Philosophical Investigations* (e.g. Hacker 1972, Minar 1998, Sluga 1996). However, there is not much literature tracing connections between these two lines of thought in Wittgenstein. My claim in this essay is that Wittgenstein's discussion of rule-following and meaning leaves us with a radically novel conception of the meaningfulness of language and that this in turn underpins his rejection of Cartesian dualism and traditional assumptions about the self that inform both realist and solipsistic positions. These connections are based largely off the first part of the *Philosophical Investigations*. In tracing these connections I hope to defend Wittgenstein's position as a tenable alternative, though it needs more elucidation and elaboration in future research.

Chapter one focuses on dismantling what I call a ‘traditional assumption about meaning’ (TA), which claims that the meanings of words and the actual physical uses of

linguistic expressions themselves are two separate and discrete entities. Wittgenstein explicitly mentions this assumption as a ‘mistake’:

The mistake we are liable to make could be expressed thus: We are looking for the use of a sign, but we look for it as though it were an object *co-existing* with the sign. (One of the reasons for this mistake is again that we are looking for a “thing corresponding to a substantive.”) (BB, p. 5)

Here, Wittgenstein treats the ‘use’ of a sign as its meaning, whereas the ‘sign’ is akin to my ‘actual physical use’ of a linguistic expression (a sign). The arguments put forth come from Wittgenstein’s rule-following considerations and are intended to provide strong reasons for pursuing a picture of meaning not dependent upon TA.

Chapter two sketches my take on Wittgenstein’s conception of meaning. The essential point of this chapter claims that meanings are not entities at all. Rather, the meaningfulness of any expression cannot be separated from the actions and circumstances surrounding an expression’s use. This point does not endorse TA, thereby avoiding the problems raised in the first chapter. An elaboration of this point consists in two main moves: (1) discussing the meaningfulness of our most primitive language, which consists of primitive gestures and sounds and (2) providing a possible way that we can use the meanings found in our more basic language to create more extensive meanings in the highly developed language we use today.

Chapter three connects the previous discussion of meaning to Wittgenstein’s famous private language argument (PLA). I draw a distinction between what I call the negative PLA and the positive PLA. The negative PLA concludes that there can be no private language if the meaning of our words is based off of TA. This means that no one can have a language that s/he in principle could only understand. However, the positive PLA concludes that an individual in isolation can develop a language, but not that this

language would be impossible for others to understand. While the conclusion of the negative PLA follows implicitly from the results of the first chapter, I offer the entire PLA here as a way of recapping our results and completing the previous discussions. The chapter then connects these discussions to the Cartesian assumption (CA) that mind and body are distinct substances and elucidates the new picture we should have of the meaning of our sensation language. This new picture rejects CA, advocating a kind of primitive notion of a human being from which the notions of mind and body are derived. Our sensation language depends on this primitive notion of a human being, even though it is an augmented refinement of our primitive language.

Chapter four takes the conclusions of the previous chapters and applies them to Wittgenstein's treatment of solipsism. After articulating the solipsist's position, I point out that this position relies on a notion of meaning and privacy based off of TA and CA that we have rejected as being untenable. Therefore, the solipsist's thesis is not genuinely meaningful. This makes the solipsism incoherent, in which case we can dismiss solipsism as a possible position.

Chapter One:

The Rule-Following Problem

A large part of the *Philosophical Investigations* is devoted to what have been labeled “The rule-following considerations”. Unfortunately, these arguments are not simple, and have generated conflicting exegeses (Kripke 1982, Baker and Hacker 1984). Wittgenstein’s style might be aesthetically pleasing and original, but it certainly does not facilitate a univocal interpretation. In any case, the arguments in this chapter focus on attacking a central traditional assumption about meaning via the rule following considerations. Generally speaking, ‘the traditional assumption about meaning’ (TA) claims that there is a sharp distinction between physical uses of linguistic expressions and the ‘meanings’ that attach to them. These are two *separate and discrete* entities, and it is the job of the philosopher to establish the relationship between them.

This chapter begins by stating the problem and some initial responses, followed in the second section by a critique of the dispositional response. The third section of this chapter attacks two substantial responses to Wittgenstein’s rule following considerations. One adopts a broadly Fregean distinction between meaning (sense) and linguistic statements while the other advocates a kind of Russellian conception of immediate acquaintance with sense-data that determines meaning. These are two highly influential ways in which meanings and linguistic expressions have been related, and so before going further I will spend some time to clearly illustrate the major assumptions of these positions that are relevant to our purposes. However, it should be noted that the aspects of the views highlighted below are not meant to apply solely to their respective fathers of thought. I mention these aspects of Frege’s and Russell’s thought to point out two highly

influential assumptions about meaning that are present in many other theories of meaning, which may or may not differ from Frege or Russell. These assumptions motivate responses to Wittgenstein's questions, but all fall short of their goal. Ultimately, the chapter concludes with what I call the negative private language argument, which claims that there can be no act of private ostensive definition that determines the meaningfulness of words. Hence, no one can 'obey a rule privately'. Since this kind of privacy is necessary for determining the meaningfulness of our words in both Frege's and Russell's theories, we must abandon them as viable accounts.

For Frege, each linguistic statement has a subject and predicate. These *may* also have their own senses and referents as well. Consider, for example, the sentence,

(1) Hilary Putnam is a mathematician.

(1) has a subject, "Hilary Putnam" and a predicate "... is a mathematician". The sense of "Hilary Putnam" is an abstract concept and this in turn fixes the referent, the actual man Hilary Putnam. The predicate plays a similar role. Its sense is an abstract functional concept that refers to an actual objective function. The former takes the senses of subjects while the latter takes actual objects as input values. The sense of (1) in its entirety is an abstract, objective *meaning* that results from the combination of the senses of the subject and predicate. The referent of (1) is a truth value obtained in an analogous relationship among the referents of the subject and predicate.¹

¹ This is an admittedly cursory summary of Frege. For example, Frege holds that the subjects and predicates of linguistic statements may each lack a sense or a referent, respectively. "Unicorn" does not have a referent because there are no actual unicorns and "Triffomite" does not have a sense because it has not been given one in our language. In the former case, any sentence using "Unicorn" as a subject would lack a truth value but still have a meaning while, in the latter case, a sentence using "Triffomite" would also lack a meaning. Furthermore, Frege's philosophy helps us deal with problems of reference, meaning, and knowledge, such as the infamous "Morning Star" and "Evening Star" problem, and many others besides. It is an admirable system on the whole, but for our purposes we shall focus on one particular aspect of it.

The main aspect of Frege's theory that chapter one's critique focuses on is its notion of abstract meanings. For Frege, these are *objective* meanings that exist in a 'third realm'. They are akin to Plato's forms in several ways. They are immaterial, timeless, and we 'grasp' them with our minds. Furthermore, they can be 'infinite objects' thanks to their non-physical natures. That being said, Frege clearly endorses TA. We have a separate, even objective 'meaning' that is distinct from the linguistic entity with which it corresponds. Understanding the meaning of a sentence requires an act of private grasping of an objective object by our minds. The first chapter attacks this vital part of Frege's thought.²

The intricacies of Russell's thoughts on meaning are far too extensive to dwell on here, if not simply for the fact that he constantly changed his views throughout his lifetime. To make things *slightly* simpler, I will focus on his logical atomism and his general theory of meaning expressed therein. Russell's logical atomism breaks down ordinary sentences, by means of logical analysis, into 'logical atoms', which are essentially sense-data of varying kinds.³ In "The Philosophy of Logical Atomism", Russell says:

The reason that I call my doctrine *logical* atomism is because the atoms that I wish to arrive at as the sort of last residue in analysis are logical atoms and not physical atoms. Some of them will be what I call 'particulars' – such things as little patches of colour or sounds, momentary things – and some of them will be predicates or relations and so on. (Russell 1985, p. 37)

² This is a woefully cursory summary of Frege, but I hesitate to dwell any further on elaborating his thought. The main point to emphasize is that meanings are objective abstract entities that we somehow grasp with our minds and that it is this process that establishes the connection between linguistic entities and meanings.

³ This is related to Russell's endorsement of a kind of radical empiricism. For Russell, "Sense-data" include direct impressions of actual relations, qualities, and even facts. All of these are parts of our sense-data, not just sense-impressions of physical objects. I only note this here to avoid confusion about Russell's stance.

Logical analysis takes our ‘complex’ ordinary sentences and breaks them up into their ‘simple’ component parts. In the final analysis, we are left with completely simple symbols whose meanings come from direct acquaintance of particular sense-data.

All analysis is only possible in regard to what is complex, and it always depends, in the last analysis, upon direct acquaintance with the objects which are the meanings of certain simple symbols. (Russell 1985, p. 54)

To understand the meanings of complex symbols, we must first understand the simple symbols of which they are composed, and the meanings of these symbols comes from direct acquaintance with particular objects of sense-experience. For example, Russell considers the word “red” to be paradigmatic of a simple symbol, and the meaning of red is known through direct acquaintance with red patches of color.

There are finer details of Russell’s logical analysis that warrant further discussion elsewhere but for our purposes the main points of his account lie in his emphasis on sense-data.⁴ Our ability to know the meanings of any symbol in our language, whether simple or complex, boils down to our ability to have direct acquaintance with relevant sense-data. This argument commits us to a Cartesian conception of the mind that divorces appearances (sense-data) from reality.⁵ The sense-data private to a particular subject determines the meaning of simple symbols, presumably by a means of some private ostensive definition. While many people adopt this influential aspect of Russell’s theory, doing so commits them to TA. The meaning of a symbol and the symbol itself are *two separate and discrete* entities. The meanings of simple symbols lie in private experience, which are determined by an act of private ostensive definition that establishes the

⁴ For example, Russell’s discussion of the cognitive aspects involved in meaning seem relevant, though tangential, for our purposes (Russell, 45-6). His elaboration of analysis, definitions, and the meanings of complex symbols (Russell 54-7) is somewhat pertinent to this essay as well, if not for the sole fact that the later Wittgenstein criticized Russell on these points (PI 46, 55, 59, 60).

⁵ John McDowell discusses Russell’s Cartesian commitments thoroughly in his article, “Singular Thought and the Extent of Inner Space”. We will see the role Cartesianism plays later in the essay.

relationship between a simple symbol and its meaning. This is the notion that the arguments below refute.

I. The Initial Problem

Mathematics is a language game extensively used by Wittgenstein to express his concerns, and also serves as a clear illustration of his problems. Consider the series:

1) 2, 4, 6, 8, ...

We have been taught to regard this as the elementary series of even integers that any primary school student (or possibly even a child before formal schooling) will master. It follows the algebraic formula “ $2n$ ” where “ n ” belongs to the set of integers. The main thrust of Wittgenstein’s rule-following sections is expressed by the question, “How does the student know how to go on *in the correct manner*?” In other words, how does the student come to know the correct meaning of “ $2n$ ”? Given this finite sequence, any number of rules could occur to the student, such as:

- a.) $2n$
- b.) $2n + (n-1)(n-2)(n-3)(n-4)$
- c.) $2n$ for all $n < 4$, and 5 for all $n \geq 4$

These are just a few of the infinitely many possible rules that could occur to the pupil.

This exact example is discussed by Wittgenstein in PI 185:

Now we get the pupil to continue a series (say +2) beyond 1000 – and he writes 1000, 1004, 1008, 1012.

We say to him: “Look what you’ve done!” – He doesn’t understand. We say: “You were meant to add *two*: look how you began the series!” – He answers: “Yes, isn’t it right? I thought that was how I was *meant* to do it.” – Or suppose he pointed to the series and said: “But I went on in the same way.” – It would now be no use to say: “But can’t you see?” – and repeat the old examples and explanations. – In such a case we might say, perhaps: It comes natural to this person to understand our order with our explanations as *we* should understand the order: “Add 2 up to 1000, 4 up to 2000, 6 up to 3000 and so on.” (PI, 185)

Perhaps the difficulty is not readily apparent. It does not take a brilliant mathematician to realize that there could be any number of algebraic formulas that would give us different numbers for the value of $n = 5$, and, moreover, this is not problematic *prima facie*. Things become problematic when we try to pin down how it is possible for us to teach the rule “ $2n$ ” to the student. When can we say that the pupil has learned the formula “ $2n$ ” and not some other strange formula like “ $2n + (n-1)(n-2)(n-3)(n-4)$ ”? Furthermore, how do *we*, as the teachers, accurately convey to the student that we *mean* “ $2n$ ” and not “ $2n + (n-1)(n-2)(n-3)(n-4)$ ”? Clearly not after we have progressed along to a certain point in the series, for we can only physically demonstrate a finite segment of the series. Conversely, we cannot say that the student knows what rule he is following after any definite value of n , such as 20, or 100, etc.

The most immediate suggestion is to offer some explanation as the following. “Look, the series is meant to be quite simple. All the pupil has to do is *add two* after *every* spot in the series. He must do the *same* thing every time.” These are certainly instructions for the student to follow, but they are merely a more verbose expression of “ $2n$ ”. This option does not offer any kind of solution because the problem cannot be addressed by simply substituting one linguistic expression of a rule for another. Even if one were to specify the new linguistic expression such that mathematical words like “add” or “two” were absent from it, we could still legitimately repeat our question at this more basic level. This strategy results in an infinite regress of substitutions of linguistic expressions for one another, and therefore lacks any explanatory power.

II. The Dispositional Response

The solution to this problem cannot be straightforward, so before diving into more complicated responses we should flesh out what Wittgenstein's opponents can appeal to and how he intends to combat them. Firstly, Wittgenstein does not restrict possible explanations to purely physical or behavioristic options. One might want to brand his conclusions as a form of behaviorism, but, as Kripke points out (Kripke, 44), he does not assume behaviorism as a premise. This gives Wittgenstein's interlocutor the ability to appeal to mental states when trying to respond to his rule following questions. Furthermore, Wittgenstein reacts against many of his earlier views expressed in the *Tractatus Logico-Philosophicus*, which implicitly adopt some form of TA. It is often helpful to view his interlocutor in the PI as an advocate of TA and, in some situations, as his younger self. Wittgenstein should be read as debating with these speakers *on their own terms*; he tries to show that they fail to adequately explain the meaningfulness of our language based on their own assumptions.

Since we have to look to something beyond a more careful description of the rule itself to help us, someone might insist "But I knew that the next number needed to be 10! I was thinking of this as I was teaching him, and that is what determined how I meant the correct next step." Perhaps this person did have the next step in mind, but she could not have had *all* of the infinitely many correct steps in mind. It is simply impossible for her to have thought of every single member of this series because there are infinitely many of them.

Wittgenstein considers a more robust version of this response, which is essentially a response that appeals to the counter-factual, "If I had been asked what number came next, I would have responded with 10." This could also be restated as a *disposition* to

behave in a certain way. The reason that the teacher *meant* to teach the pupil the rule, “ $2n$ ” is that she was *disposed* to continue the series with “10, 12, 14, etc.” Furthermore, if she had been asked about any particular term in the series, she would have responded with the correct response because she has the necessary kind of disposition that accords with the rule. Kripke discusses the faults of the dispositional response at length in the second part of his celebrated essay, “Wittgenstein on Rules and Private Language”, and the criticisms I put forth against dispositionalism will mirror Kripke’s.

The first reason for abandoning dispositionalism is that our dispositions are finite, which is ironic given that the move towards dispositionalism was intended to overcome problems about finitude, as Kripke points out:

The dispositional theory attempts to avoid the problem of finiteness of my actual past performance by appealing to a disposition. But in doing so, it ignores an obvious fact: not only my actual performance, but also the totality of my dispositions is finite. (Kripke 1982, p. 26)

In the case of this series, it is easy to give the correct answer for any reasonable term in the series, say 15,000,000 or some smaller manageable number. We simply multiply these by two to get the correct answer. However, it seems less likely that I have a disposition to give the correct answer when asked to give the 235,986,857,100,883,345th term in the series. That really just seems like a huge mess of numbers, and while I could calculate the right answer, it will take some time and work (although a modest amount of each). More importantly, the main force of this criticism stems from the fact that there are simply numbers that are too large for us to have a disposition towards because we could never actually form a disposition towards them. There are infinitely many numbers, and we exist for only a finite amount of time. This means that we can only have a finite number of dispositions. Therefore, there are infinitely many numbers that we will never

be able to have a disposition towards in our life times simply because we will die before we can form any disposition towards them. How, then, are we to determine that the student is not following some rule such as: “ $2n$ for all n physically conceivable, and 7 for all n too large for us to form a disposition towards” rather than, “ $2n$ ”?

This initial criticism is admittedly unappealing and seems to miss the value of the dispositional theory. “If we could overcome our physical limitations and actually grasp the number desired,” the dispositionalist might frustratingly reply, “we would respond in accordance with our dispositions and therefore give the correct term for the series.” This brings us to a more important point: since *these* dispositions are now *theoretical* dispositions that we *would* have under ideal conditions, how do we determine what we would be disposed to do? To say that our dispositions under ideal circumstances would be those that accord with the intended series would be circular because we are trying to figure out how anything can accord with the intended series at all. An appeal to ideal circumstances cannot overcome this problem with dispositionalism.

Another pressing worry with the dispositional account claims that this account lacks the normative component necessary for correct rule following. In point of fact, we are sometimes disposed to follow rules incorrectly. “Horse-looking” cows have become a popular example to illustrate this point (see Boghossian, 1989). Presumably, I have correctly learned what objects are to be considered “cows” and the others that are “horses”. But if I found myself in a sufficiently vague scenario, say a dark, foggy night and at a distance from a fairly abnormally shaped cow, I might be *disposed* to call the cow a horse. But this does not make it the case that “horse” means “either a horse *or* a cow”, though I am disposed to use it as such. More generally, the force of this objection

boils down to the fact that just because I *feel* that the next term in the series is 'x' does not make it the case that 'x' is the *correct* term. Just because I happen to be disposed in such a manner to react in such a way does not mean that this is the right way to respond.

I find these problems devastating for an adequate dispositionalist response to our problem. By adopting TA, we have to bring meaning and the actual uses of linguistic entities into accord somehow, but a dispositional theory faces two difficult objections: (1) since the meaning of "2n" applies to infinitely many numbers, our dispositions fail to uniquely determine a meaning for any given expression and (2) since our dispositions are based on merely what we *feel* is right, we quickly run into extreme problems of relativism that prohibit us from correcting anyone's application of a word. If the student continues the series "2n" past $n = 100$ and she starts writing, "204, 208, 212, etc." we cannot legitimately reprimand her. Instead, we can only say, "Well, she was disposed to do this, so it must be right." This arrests any possibility of teaching in the first place, because all of the students *can only* ace their tests!

III. Frege's and Russell's Responses

At this point, we come to more weighty responses to Wittgenstein's challenge. I will focus on Frege's assumptions first because they pose less of a challenge to Wittgenstein. Some of the problems above are avoided by Frege. Frege postulated a mental faculty that we possess that allows us to 'grasp' objective meaning facts, or something of the kind. For Frege, the meanings of propositions are not subjective feelings; they are actual objective entities that we 'grasp' with our mental faculties. The objective nature of these entities overcomes several problems. Firstly, they could be

‘infinite’, in a sense, thereby avoiding a serious problem facing the dispositional account. The mysterious nature of our mental faculty accounts for our ability to grasp such a ‘meaning’, which can supposedly determine how we are to use linguistic expressions in infinitely many cases. Furthermore, since these are objective entities that we grasp through a faculty, we can sense meanings in this higher ‘third realm’ in a similar way in which we sense the external world. We are given a picture of how this sixth sense might be able to pin down the meaning we seek.

Appealing to Frege in this manner only moves the problem about rule following to another level. How do we know when someone has ‘grasped’ the intended meaning? The rule “ $2n + (n-1)(n-2)(n-3)(n-4)$ ” exists in Frege’s third realm too, and we need a way of establishing how to attribute meaning to a student whose mind is still grasping at straws. Kripke succinctly discusses this kind of Platonist route towards the end of his second chapter:

But the ultimate sceptical problem cannot be evaded, and it arises precisely in the question how the existence in my mind of any mental entity or idea can *constitute* ‘grasping’ any particular sense rather than another. The idea in my mind is a finite object: can it not be interpreted as determining a quus function, rather than a plus function? Of course there may be another idea in my mind, which is supposed to constitute its act of *assigning* a particular interpretation to the first idea; but then the problem obviously arises again at this new level. (Kripke 1982, p. 54)

While these meaning entities might be ‘infinite’, our minds certainly are not. So the question arises again: how do we determine that my mental act of grasping constitutes “ $2n$ ” and not “ $2n + (n-1)(n-2)(n-3)(n-4)$ ”? I have a finite idea in my mind that is supposed to correspond in some fashion to an objective meaning entity, but its finiteness allows for infinitely many possible meaning entities to be related to it. Essentially, Frege has to rely on a notion of private ostensive definition, which we will discuss at length at the end of this chapter. This dependence upon the workings of a private mental realm destroys any

advances a theory of meaning might have through an appeal to objective meaning entities. Frege was a brilliant thinker on his own, but this particular aspect of his thought cannot help us with our problem.

Russell's response presents a greater challenge to Wittgenstein, but ultimately relies on an analogous conception of private ostensive definition. The essence of Russell's stance lies in the fact that the meaning of an expression boils down to a sensation in the mind of the learner. In our example of "2n", the student becomes directly acquainted with a sense-datum, a particular private experience, that *constitutes the meaning* of "2n", presumably through familiar methods of training and repetition. This response does have some explanatory power and initial plausibility. We do seem to have some kind of sensation associated with those elementary epiphanies we had when learning arithmetic, and continue to have other sensations as we solve new puzzles, study new languages, or learn things in general.

Unfortunately, private sensations do not hold water as a helpful explanation for very long. In section 160, Wittgenstein discusses a similar problem with knowing when someone is 'reading':

But imagine the following case: We give someone who can read fluently a text that he never saw before. He reads it to us—but with the sensation of saying something he has learnt by heart (this might be the effect of some drug). Should we say in such a case that he was not really reading the passage? Should we here allow his sensations to count as the criterion for his reading or not reading? (PI, 160)

Applying this criticism to our case, grant that we have an idea of what this sensation is, and further suppose that we (harmlessly) induce the sensation in a child who is learning arithmetic. We ask them to perform a simple sum, say "49 +21", and the child shouts, "Oh I know how to do that! The answer is 72!" Are we to say that the student has learned

how to add or not? Despite having the sensation supposedly required for meaning addition, it is clear that the student does not understand how to add.

There are other simple considerations that make this response unappealing. If meaning is supposed to be this private mental sensation accompanying linguistic expressions, how are we to know that each of us means the same thing when we communicate?⁶ These sensations are in principle inaccessible to anyone else. An argument from analogy gets us a probabilistic certainty *at best*. Furthermore, the whole qualitative character of this sensation is enigmatic. Kripke illustrates this problem well:

Attend to what happened when I first learned to add. First, there may or may not have been a specifiable time, probably in my childhood, at which I suddenly felt (*Eureka!*) that I had grasped the rule for addition. If there was not, it is very hard to see in what the suppositious special experience of my learning to add consisted. Even if there was a particular time at which I could have shouted "*Eureka!*" – surely the exceptional case – in what did the attendant experience consist? Probably consideration of a few particular cases and thought – "Now I've got it!" – or the like. Could just *this* be the content of an experience of 'meaning addition'? How would it have been different if I had meant *quus*?⁷ (Kripke 1982, p. 44-5)

What can we really say about this sensation? More importantly, how would it differ from the one we would get when we mean to follow " $2n + (n-1)(n-2)(n-3)(n-4)$ "? There is not a readily perceivable way to distinguish between the two. And suppose that there were a way; when are we to say that this sensation came to us? When I wrote out the n^{th} term in

⁶ Evidently, there was a time when Russell did not find this objection problematic. Indeed, it seems that he thought that if we meant the same things by our words, we "...would be unable to tell people at home what you had seen in foreign parts [of the world]." (Russell , p. 56) He recognizes that no one has the same exact sense-data experience, and so, according to his theory, no one means the exact same things with their words. However, he thinks this is necessary for us to successfully communicate. Needless to say, I find the possibility that we might never mean the same things by our words extremely problematic. While Russell's theory demands more attention, it is too tangential presently. In any case, his theory does not seem to meet Wittgenstein's overarching challenges.

⁷ Kripke uses the word "*quus*" here to denote the function he introduces as "quaddition" in his argument. Unlike "addition", "quaddition" is the arithmetical function that dictates: "For any pair of numbers $x, y < 57$, perform usual addition to find their 'quum'. If either x or y is equal to or over 57, respond with "5" for their quum." This is just a different example of a rule to which all of our behaviors have conformed, similar to " $2n$ for all $n \leq 4$ and 5 for all $n > 4$ " in our example.

the series? This response faces a host of perplexing questions and it does not seem to be equipped to deal with them.

These initial considerations are perhaps not fair enough to Russell and advocates of his assumption. There are two related objections to consider. The first argues that anyone can differentiate between the sensations of grasping the meaning of “ $2n$ ” and grasping the meaning of “ $2n + (n-1)(n-2)(n-3)(n-4)$ ”. These sensations are as different as the differences between the sensations of blue and red. The student has direct acquaintance and, moreover, *infallible access* to her sense-data, and so it does not seem possible that she might misidentify her sensation for an incorrect ones. The second concedes that there might not be an immediate sensation for “ $2n$ ” with which we become directly acquainted, but that is because “ $2n$ ” is a *complex* expression. The meanings of our simple symbols do lie in such an acquaintance. The meanings of “blue” and “red” lie in the familiar experience we all have with patches of blueness and redness. The meanings of more complex expressions, such as “ $2n$ ”, are constructed out of these simple symbols and various logical relations that obtain between them.

These responses bring us to some crucial points of the rule-following considerations. Concerning the first, we are left with at least two substantial problems even if we grant the student infallible access to her sensations. Firstly, since our sensations are finite, it is not clear how the *experience* of “2, 4, 6, 8 ...” is supposed to manifest the ‘sensation’ “ $2n$ ” rather than “ $2n + (n-1)(n-2)(n-3)(n-4)$ ”. If we consider the series “ $2n$ ”, we see that we can only experience a finite segment of the series. Presumably after some finite term in the series, we privately ostensively define the meaning of “ $2n$ ”. But what *quality* of our sensation determines that we now mean “ $2n$ ” and not “ $2n + (n-$

1)(n-2)(n-3)(n-4)”? We *only experience* seeing the numbers “2, 4, 6, 8, ...”, and so claiming that this experience gives us a sensation of “2n” rather than “2n + (n-1)(n-2)(n-3)(n-4)” gives an unjustified preference for the former over the latter. In this respect, this case is *not* like the cases of blue and red, for our sensations are equivalent.⁸

Regarding the second objection, restricting our considerations to cases of blue and red does not dodge our difficulties. There are at least two strong objections facing this move. We can concede that there is a fundamental difference between the sensations of red and blue. However, a worry arises even when we consider the nature of the basic sensation of red closely. What, exactly, *is* the sensation of red? Does it have a particular shape, say a square, or is it irregularly shaped?

Ask yourself: what *shape* must the sample of the colour green be? Should it be rectangular? Or would it then be the sample of a green rectangle? – Should it be ‘irregular’ in shape? And what is to prevent us then from regarding it – that is, from using it – only as a sample of irregularity of shape? (PI, 73)

Why do we not take our particular sensation of red to mean “red squares” or “irregular shapes”? Any particular sensation we have is *not sufficient* for establishing the meaning of “red” because the sensation could be interpreted to stand for various words with dramatically different meanings.⁹

⁸ Appealing to radical empiricism will not help in this case either. Let us grant the plausibility that we actually experience relations between objects and, less controversially, that we experience their qualities. The question still remains: what is it about viewing this finite segment of the series that gives you the actual *sensation* of “2n” and not “2n + (n-1)(n-2)(n-3)(n-4)”? I might be able to perceive the relation “x is taller than y”, but I do not perceive the infinitely many numbers of this series and any pattern or relationship between the ones I do see could be interpreted under either rule. It is not clear how we could demonstrate how one rule manifests itself to our senses over the other without begging the question.

⁹ The problem above leaves us with a problem concerning the indeterminacy of private ostensive definition. Our particular sensations cannot determine the precise meaning we are after. A similar point is made against referentialism in *The Blue Book* with Wittgenstein’s example of the word “tove”. (BB, p. 2) Referentialism holds that the meanings of words are the particular *external* objects for which they stand. The “tove” example is supposed to establish a fatal indeterminacy of reference plaguing this account as well. Translating his example to our present case, it is unclear how we could definitively ostensively refer to the color red, or at least we cannot be sure that another picks out the object we intend them to. The color red inheres in various objects, and so it appears impossible to definitively ostensively point to the color red and not books, tables, shapes, quantities, and so on.

Now let us suppose that the objection above fails somehow, so we grant that we have access to the sensation presumably required for successful private ostensive definition.¹⁰ Despite this concession, the meaning of “red” is surely not exhausted by *this particular* sense-impression of redness. The meaning of “red” has a certain generality and limitless applicability that goes beyond any immediate sense impression you might have. This requires a method by which we can determine *correct* uses of the newly defined word in the future. The most natural attempt to formulate such a method relies on appealing to past sensations. When I see a new patch of color, I know whether or not this color is “red” that I privately defined it in the past because it either is or is not *the same* color I previously had. Since we know what our sensations are and are allowed to distinguish between them coherently, picking out the same one on any given instance should not be a problem.

Even with coherent access to sensations, it looks like this account will fail. Consider being presented with a patch the color we call ‘red’ and how we might privately ostensively define a meaning corresponding to it.

A definition surely serves to establish the meaning of a sign. – Well, that is done precisely by the concentrating of my attention; for in this way I impress on myself the connexion between the sign and the sensation. – But “I impress it on myself” can only mean: this process brings it about that I remember the connexion *right* in the future. But in the present case I have no criterion of correctness. One would like to say: whatever is going to seem right to me is right. And that only means that here we can’t talk about ‘right’. (PI, 258)

The first problem with private ostensive definition is that we have no criterion of correctness. Whatever sensations we have in the future will mean “red” because we choose them to. But the problem with a private language extends beyond this resulting

¹⁰ We should reconsider Frege’s position as well at this point, because the arguments below apply perfectly analogously to the supposed idea in our minds that determines an objective meaning entity. Frege relied just as heavily on the notion of private ostensive definition as Russell did; they just thought ‘meanings’ were different kinds of entities.

relativism about meaning. There is actually no way for us *to follow* our own “private rules” because we are simply using the words *whimsically*. This gives us a horrible problem of indeterminacy of meaning: *we* do not even know what the correct use of any word is.

Still, we might want to appeal to some ability to ‘subjectively justify’ the uses of our words. Surely we know the procedure we use to determine what word to call a given sensation: this consists in comparing the given sensation against our previous ones. We establish a method of comparison (i.e. a private rule that we follow) and then apply the method in future uses.

There are a number of problems with ‘subjective justification’. Firstly, the method by which we compare previous sensations to our current ones lacks a criterion of correctness as well. For example, if we were to try to justify our use of the word “red” upon perceiving a certain color, we might try to appeal to whether or not this sensation is “the same” as the past sensation we used to privately ostensively define a word. However, determining whether or not this sensation is “the same” as our past sensation *is completely up to us*. We are not “following a rule” in any sense because there are no constraints on future “correct” uses of a word:

Don’t always think that you read off what you say from the facts; that you portray these in words according to rules. For even so you would have to apply the rule in the particular case without guidance. (PI, 292)

It is clear that we might postulate some second order rule governing the application of another rule, but our worries simply arise again at this level. Appealing to higher order rules only results in an infinite regress of rules where the problem repeats itself at each new level.

All of the above considerations constitute what I call Wittgenstein's negative private language argument. Appealing to a notion of private ostensive definition cannot determine the meanings of our words. We are left with a large mess about rules in the wake of this conclusion, but we have been working towards this for a purpose. We all have a deeply rooted picture that our language is governed by grammatical rules, and this idea extends beyond the concept of grammar that we all have from primary school. There are generally supposed to be particular constraints on the correct application of words, "Green", "Table", "And", and all of the other words in our language. If a child calls something that is red "Green", we chide them and realize they have not learned the rules governing the word "Green". Similar things happen for all of our words, but the considerations above cast doubt upon our natural conceptions of language acquisition. One cannot construct a meaningful private language in the manner considered above, or at least one that relies on these Fregean or Russellian notions of private ostensive definition and rule-following. Furthermore, we see the devastating affects resulting from adopting the traditional assumption about meaning (TA). Both Russell and Frege endorse TA, and they run into dramatic problems when they try to establish how meanings can be brought together with our uses of linguistic expressions.

However, our language is clearly meaningful. I take this as a brute fact and I do not see anyway of denying this statement without being falling into a bizarre realm of problems and incoherencies. If our entire language is *not* meaningful, what does that thesis really assert? It is supposedly trying to tell us something... but it cannot? The thesis is not even straightforwardly self-contradictory because all contradictions are at least *meaningful*. We can understand what contradictions say, even though their logical

form makes them false necessarily. This thesis cannot even get off the ground. Therefore, we must reject it for failing to tell us *anything at all*. With that settled, we can begin to describe how language is actually meaningful.

Chapter Two:

Uprooting Old Assumptions and Planting New Meaning

This chapter will sketch a picture of how our language can still be meaningful without the traditional assumption about meaning (TA) criticized in chapter one, respecting Wittgenstein's stubborn refusal to give a definite *theory* about meaning. While this rough picture is intended to avoid all of the problems Wittgenstein raises about TA, I present it in a somewhat infantile stage. My only purpose in doing so is to present a new way of looking at meaning that could be developed further elsewhere and does not face any immediate problems about rule-following. I do not intend to profess the infallibility of this picture, but merely that it is a tenable alternative. The rest of this essay should be read as a meditation on how this alternative, if it can bear fruit, could help us address the problem of other minds, solipsism, and ultimately the self. To this end, I am indebted to Saul Kripke, Hilary Putnam, John McDowell, P.M.S. Hacker, and G.P. Baker for their elucidating secondary commentary, and my debt will be acknowledged when necessary.

Before discussing his rejection of TA, I will clarify how Wittgenstein avoids the familiar labels within contemporary metaphysics of Realist or Anti-realist.¹¹ This will give us a deeper appreciation for how we can maintain a coherent philosophical position despite rejecting a seemingly necessary assumption. The next section elucidates the role of meaning in our most primitive language games, and the final section discusses a possible way in which our current languages can be seen as various extensions of this

¹¹ Broadly speaking, a Realist stance holds that there is a readymade objective or mind independent world that we have direct access to via our language and thought. Anti-realism claims that we do not have access to reality as it is in itself, as it is objectively. Instead, we only have access to experience tainted by the permeated presence of our thought or language, depending on the brand of Anti-realism one advocates.

primitive language without reducing our more complex meanings merely to those found in our primitive language.

I. A Revolutionary Rejection

At first it seems that Wittgenstein, in the spectrum of theories between Realism and Anti-realism, holds a form of Anti-realism, while his exact place in the continuum may be arguable. Here are a few quick quotes that support this Anti-realist reading:

Essence is expressed by grammar. (PI, 371)

The proof doesn't *explore* the essence of the two figures, but it does express what I am going to count as belonging to the essence of the figures from now on. – I deposit what belongs to the essence among the paradigms of language.

The mathematician creates *essences*. (RFM, 32)

Grammar tells what kind of object anything is. (PI, 373)

How do I know that this colour is red? – It would be an answer to say: “I have learnt English”. (PI, 381)

These maxims smack of Anti-realism. We can only conceive of the essences of things through our language. There are no ‘kinds of objects’ that exist independent of our grammar. Answers to classic metaphysical problems about properties dissolve simply by understanding the English language. Aside from these specific quotes, Wittgenstein’s notion of language games is thoroughly Anti-realistic. None of our language games depict or correspond with “Reality” in any sense. All we have are various games that are not even ‘true’ in any metaphysically superlative sense. When we try to ask philosophical questions about the nature of reality, we only bump into the limits of our language and end up asking meaningless questions – making illegitimate moves in our language games.

Despite these strands in Wittgenstein's thought that scream of Anti-realism, there are several passages that seem to indicate sympathy towards a Realist view. Here are a few:

(*Tractatus Logico-Philosophicus*, 4.5): "The general form of a proposition is: This is how things are." – That is the kind of proposition that one repeats to oneself countless times. One thinks that one is tracing the outline of the thing's nature over and over again, and one is merely tracing round the frame through which we look *at it*. (PI, 114, emphasis added)

The aspects of things that are most important for us are hidden because of their simplicity and familiarity. (One is unable to notice something – because it is always before one's eyes.) [...] And this means: we fail to be struck by what, once seen, is most striking and most powerful. (PI, 129)

These passages retain the elements of Anti-realism expressed above. However, they have a remotely Kantian flavor, and even seem to hearken back to the infamous seventh proposition of the *Tractatus*: "What we cannot speak of we must pass over in silence." Wittgenstein certainly rejects all talk about 'things-in-themselves' as non-sense, but here he is not saying that we cannot know things as they *really* are. Rather, he seems to advocate what might be called a *quietist* Realism. I believe that Charles Sanders Peirce hints at a helpful version of this position in his article "How to Make Our Ideas Clear":

But it may be said that this view is directly opposed to the abstract definition which we have given of reality, inasmuch as it makes the characters of the real depend on what is ultimately thought about them. But the answer to this is that, on the one hand, reality is independent, not necessarily of thought in general, but only of what you or I or any finite number of men may think about it; and that, on the other hand, though the object of the final opinion depends on what that opinion is, yet what that opinion is does not depend on what you or I or any man thinks. (Peirce 1995, p. 38-9)

I quote Peirce with caution, because Wittgenstein certainly has many points of disagreement with him. Peirce's deference to the scientific method, his theory of signs, and his overall commitment to a more robust scientific Realism are fairly un-Wittgensteinian. Let us therefore confine ourselves to this quote alone: we see that *ultimately* the objects of our opinions are determined by what opinions we have (i.e. Grammar tells us what kind of object anything is), and so, when we try to *say* anything,

we fall short of describing reality “as it is in itself.” On the other hand, *there is an independent reality*, and it forces itself upon us in ways that constrain our ability to arbitrarily create meaningful language solely on the basis of our subjective whims. From Wittgenstein’s angle, the reality Peirce speaks of is what is *most familiar* to us because it is always in front of us. While we can trace and retrace our frame, we are always *looking at something*, but whenever we try to get outside of our language to talk about it we fall into metaphysical non-sense.

Both of these readings have their merits, but ultimately I think Wittgenstein would reject the label of either Realist or Anti-realist. Hilary Putnam endorses this reading of Wittgenstein as well, and brings out helpful distinctions between Wittgenstein and Kant:

For Wittgenstein, the negation of a pseudo-proposition is a pseudo-proposition; the negation of nonsense is nonsense. If we are persuaded that it is *unintelligible* to say “We sometimes succeed in describing reality as it is in itself”, then we should realize that it is equally unintelligible to say “We never succeed in describing reality as it is in itself”, [...] In fact, one might say that it is characteristic of Wittgenstein to try to show us that when philosophers say that we *can’t* do something, say that something is *impossible*, typically the thing that they tell us it is impossible to do is a nonsense thing, an unintelligible thing; that the philosopher, as it were, seems to be telling us of an Impotence, in the way the physicist tells us of an Impotence when he says “You can’t build a perpetual motion machine” or of a barrier we can’t cross, but it turns out on examination that the barrier is a mirage, or even less than a mirage – that it is chimerical. We can learn and change and invent languages, and in them we can state truths; that *is* describing reality. If you say, “Yes, but it is not describing reality as it is in itself”, you are saying *nothing*. (Putnam 1992, p. 39-40)

Wittgenstein does not accept Kant’s distinction between phenomena and noumena because all talk of noumena is vacuous – it is not a meaningful move in our language game. However, this does *not* commit us to any typical brand of Anti-realism.

Traditionally, Anti-realists want to say that our language or thought *gets between* us and ‘reality’, and that we can never *truly* grasp things *as they really are*. But this thesis is just as nonsensical as saying we *can* describe reality as it is in itself. As Wittgenstein might

say, the negation of a pseudo-thesis is a pseudo-thesis, and so Anti-realism should be abandoned just as readily as Realism.

Some might find this move perplexing because they believe that Realist and Anti-realist positions exhaust all possible positions that one can hold. This belief relies on the notion that language and reality are separate from each other and therefore require some kind of medium or bridge to bring them into harmony together. Of course, this is a fundamental assumption that both Realists and Anti-realists share, but it is not altogether clear that we must be committed to it. In their book, “Scepticism, Rules, and Language”, G. P. Baker and P. M. S. Hacker state the problematic assumption thusly:

How then is language connected to reality? The question continues stubbornly to bewilder us. And when Wittgenstein answers bluntly: ‘There is *no* connection between language and reality’, we react with astonishment. [...] The ground is cut out from under these deep philosophical misunderstandings [...] once it is recognized that they are not viable in the light of Wittgenstein’s repudiation of the conception of language and reality as two discrete, self-contained realms of being. If the gulf between language and reality is illusory, then of course no bridge can span it. The statement that there is no connection between language and reality is not the affirmation of an antithesis to the thesis that they are connected by some mysterious mental or metaphysical bridging apparatus. It is rather a denial that there is room for any connection, for there is no gulf to span. The question ‘How is language connected to reality?’ is what is amiss. For it rests firmly on a misconceived picture. (Baker and Hacker 1984, p. 134-6)

But what are we left with if we accept that this traditional assumption is misleading? To answer this question adequately, we must first reevaluate our conception of meaning and rules. To do this, I will restate the traditional positions that I take Wittgenstein to be attacking in his rule-following considerations, and then sketch the reorientation we must do to see things aright.

The discussions in chapter one attack what I called the traditional assumption about meaning (TA). I take this position to hold that there are *two separate and discrete entities* involved when learning the meaning of our words: the uses of symbols (and the symbols themselves) and the ‘meaning’ that we attach to them. This is a natural

assumption prevalent within the philosophy of language, and indeed it seems to be something that is plausible before deep philosophical reflection. Anyone can make symbols: we draw things in the sand, sketch meaningless doodles out of boredom in our notebooks, and put to paper various other characters that have no meaning within our language. But the words in our language *do* have meaning, and presumably *understanding* this meaning requires that we ‘grasp’ the correct *rules* that normatively guide us in the future uses of our words. In order to communicate successfully, we must all come to the same understanding of the meaning of our words, and this is done in virtue of the fact that we have all grasped *the same* rules governing our language. Doing so transforms the dead symbols we produce into bearers of meaning that can be used for communication.

Wittgenstein explicitly rejects this picture. Indeed, a great deal of the PI raises deep problems for this picture, which chapter one summarizes. While I am not concerned with defending Kripke’s well-known interpretation of Wittgenstein, I feel that Kripke has been attacked unnecessarily for labeling these problems as ‘sceptical’ problems. This label is innocuous, for Kripke only intends to point out that Wittgenstein thinks there are deep problems with traditional conceptions of meaning. This is one aspect of his interpretation that I feel safe in endorsing, primarily because it seems univocally agreed upon, appearances notwithstanding. While I do not accept Kripke’s ultimate solution, it is not clear that his initial approach to the solution dramatically diverges from

Wittgenstein’s intentions:

... his [Wittgenstein’s] solution to his own sceptical problem begins by agreeing with the sceptics that there is no ‘superlative fact’ (S 192) *about my mind* that constitutes my meaning addition by ‘plus’ and determines in advance what I should do to accord with this meaning. [...] A *sceptical* solution of a sceptical philosophical problem begins on the contrary by conceding that the sceptic’s negative assertions are unanswerable.

Nevertheless our ordinary practice or belief is justified because – contrary appearances notwithstanding – *it need not require the justification the sceptic has shown to be untenable.* (Kripke 1982, p. 65-6, my emphasis)

It is certainly true that Wittgenstein does not think that there is any fact *about my mind* that determines what rule I am following, and moreover he thinks that there is no way to answer the problems he raises for TA. Instead, we simply need to realize that we do not need what the sceptic has shown to be impossible; we can reject the assumptions that brought us to these problems in the first place. The move Wittgenstein makes is similar to his move within the debate in metaphysics between Realists and Anti-realists: we do not need to accept the fundamental premise that both sides share. Unlike where we left off with that debate, Wittgenstein paints us a more substantive picture here, which I will now explicate.

II. Primitive “Language” and Thought Experiments

The best initial approach to this picture begins with a thought experiment. Imagine a group of people millions of years ago who have not yet developed any sort of linguistic practices remotely similar to our own. The only modes of communication they have recourse to are grunts, crying, screams, playful or suggestive gestures, and other familiar things of this nature. I am not a student of linguistics or anthropology, but it seems safe to say that these kinds of things were present before humans developed the beginnings of what we call “language” today. What are we to make of these gestures? It is immediately clear that they are all kinds of *actions*, and they serve a definite role in the community. Without the ability to scream, it becomes difficult to warn the group of danger. Without the ability to make eye contact with a male or female in a certain way,

the courting process becomes difficult and filled with ambiguity. Without the ability to cry, members lose the ability to draw attention to themselves when they are in distress.

Furthermore, there does not have to be a regular connection between the sounds that these people make and their actions. This can happen in two ways: (1) the same sounds can be used for different meanings and (2) different sounds can have the same meaning. Regarding (1), screaming, grunting, etc, serve multiple purposes that vary depending on the situation in which they occur. For example, grunting can be used to stop some one from annoying you, when pushing people out of the way, or a number of other things. The exact same sound could be uttered on each instance, while their meaning differs.

Conversely, different sounds could be used for the same function. Imagine the following scenario within this community. A dead animal has just been brought back from a hunting expedition and the group has finished preparing the meat. It has been a while since any of them have eaten and, since the group is fairly large, a crowd forms around the prepared food, with the members pushing and shoving each other to get their own piece. On the outskirts of this scene is a particularly hungry male, who initially tries to “grunt” his way to the food. No holes open up, and after a while the male starts screaming. These screams have the same meaning as the grunts, although they are more effective due to their increased volume. I do not intend to mean that the male gets *annoyed* or some such thing because of his peers’ unresponsiveness; *that* would cause the meanings of the scream and the grunt to diverge. The thought experiment supposes that he does not get annoyed, merely persisting in his request to partake in the meal in both cases. He does not want to convey some newly spawned frustration; he just wants to eat.

We can conclude from these thought experiments that there does not *have to be* a regular connection between the sounds and gestures of the primitive group and their actions. Wittgenstein gives another important thought experiment to illustrate this point:

Let us imagine that the people in that country carried on the usual human activities and in the course of them employed, apparently, an articulate language. If we watch their behavior we find it intelligible, it seems 'logical'. But when we try to learn their language we find it impossible to do so. For there is no regular connexion between what they say, the sounds they make, and their actions; *but still these sounds are not superfluous*, for if we gag one of the people, it has the same consequences as with us; without the sounds their actions fall into confusion – as I feel like putting it. (PI, 207, my emphasis)

The society Wittgenstein considers is not exactly the same as the primitive group we have been considering, but there are crucial similarities between them. In both, various sounds can be *used meaningfully* for various activities. The gagging statement is best understood in reference to the primitive group of people. Imagine gagging one of the members in the primitive group. This member loses the ability to grunt, scream, or make any kind of sounds that she might use to communicate. In this way, the member becomes unable to function normally in the group; her actions become confusing to her peers without the ability to make these sounds.

I find thinking about a primitive society fruitful for many reasons. Firstly, we acquire a better understanding of how we can reject the traditional separation between words and meaning that underlies TA. Since there are no clear regularities to speak of, we can see how these gestures are *completely meaningful* despite their inability to conform to any set of 'rules'. For each sound of this primitive language, the meaning of the expression is *completely bound up* with both the action of the speaker and her surroundings. *This* is why you cannot coherently separate the meaning of an expression from the instance of the expression itself. Wittgenstein has several passages that help flesh out this account:

When I think in language, there aren't 'meanings' going through my mind in addition to the verbal expressions: the language is itself the vehicle of thought. (PI, 329)

We say, for instance, to someone who uses a sign unknown to us: "If by 'x!2' you mean x^2 , then you get *this* value for y, if you mean $2x$, *that* one." – Now ask yourself: how does one *mean* the one thing or the other by "x!2"?

That will be how meaning it can determine the steps in advance. (PI, 190)

It strikes us as if something else, something over and above the *use* of the word "all", must have changed if 'fa' is no longer to follow from '(x).fx'; something attaching to the word itself.

Isn't this like saying: "If this man were to act differently, his character would have to be different". Now this may mean something in some cases and not in others. We say "behavior flows from character" and that is how use flows from meaning. (RFM, 13)

PI 329 directly expresses Wittgenstein's rejection of TA. Verbal expressions cannot be separated from their meaning. PI 190 leaves us with a question, whose answer is that we need to *do something* in order to 'mean' anything by "x!2".¹² You have to start giving demonstrations of correct applications of "x!2" in order to show its meaning. The passage from RFM rejects TA again, and follows this with a helpful metaphor. Here is a useful example that blends the comparison. Imagine meeting a husband who tells you: "I love my wife." This statement tells us two things relative to our investigation: it is a claim about his character and it is a meaningful use of the verb 'to love'. Suppose further that we observe this man in his interactions with his wife and find out that he continually neglects her, ignores her concerns, and beats her. There is an obvious initial response we would give to such observations. We would say that his character has changed; he is no longer a man with the loving character that we thought he had. But what do we say if he *sincerely* insists that he loves his wife? The *use* of his expression and the conditions surrounding it *change the meaning of the word*. He does not mean the same thing that we

¹² The following examples rely on an intuitive notion of 'rules' for the time being. A more thorough description of the role rules play in our language is given below. These examples are simply meant to develop a preliminary understanding of Wittgenstein's picture.

do by the verb ‘to love’. This serves as a good initial demonstration of the relationship between meaning and use.

With these examples in mind, we can address the problem of rule-following more thoroughly. Turning back to the primitive group, Wittgenstein’s interlocutor might want to protest, as she does elsewhere, “You take the easy way out! You have not told us what *justifies* the use of these screams and these cries of joy in the primitive group. Why should we accept *these* uses of *these* noises and gestures as *correct* moves in the primitive language game?” Indeed, Wittgenstein does initially seem to take an easy route when he says:

What we have rather to do is to *accept* the everyday language-game, and to note *false* accounts of the matter *as* false. The primitive language-game which children are taught needs no justification; attempts at justification need to be rejected. (PI, p. 200)

This terse response might seem inadequate because it forces us to accept some kind of brute facts regarding the meaning of our primitive modes of communication. However, the better response to these questions points out that the attempt to justify our primitive language implicitly adopts TA, which we have already rejected. Another thought experiment will help show where the questioner went wrong.

Suppose I am at a train station and someone starts screaming, falling to the floor writhing in pain. I immediately run over to ascertain the scenario. There is no hesitation or doubt in my mind about what is happening. The first thing to notice is that *I do not need any training* to react in the way I do.¹³ No kind of teaching is necessary for being able to respond to such primitive sounds and gestures. We can see this in infants too.

¹³ In the quote above, Wittgenstein says that the primitive language game that children *are taught* needs no justification. The games I take Wittgenstein to have in mind include some of the most basic forms of, for example, the English language, such as object recognition, simple counting games, and so on. However, I do not consider these games to be *the most basic* language games we use, and while we have not yet considered these more complicated examples, I think that Wittgenstein’s responses are even stronger for a fundamental language game that does not require teaching.

Scream at an infant and it cries. Lovingly hold it and laugh with it and it smiles back. There is something fundamental about these sounds. They incite instinctual reactions in all humans (and other animals as well) at all ages, and *we do not need them to be taught to us*.

With this in mind, we can address the question raised by the interlocutor. Why are we justified in responding the way we do? How is the *correct* response determined? These questions implicitly assume that there is a clear distinction between a rule that dictates what we have to do and the actual actions that are either in accord or conflict with the rule. Only by accepting this dichotomy do we fall into all of the problems discussed in chapter one. But our rejection of this assumption does not leave us in the dark. We have a natural and perhaps surprisingly intuitive picture before us when we think about primitive language, where the actions and surrounding circumstances are entirely bound up with meaning. Furthermore, since there is no uniformly regular connection between the sounds the primitive people make and their actions, it becomes incoherent even within the traditional framework to separate the rule from its corresponding action. This forces us to revise our conception of the nature of a 'rule'. Rules were deemed necessary for *normatively determining the meaning of an expression*, but we have been forced to reject the idea that rules are generalized theoretical abstractions from their instantiations that we somehow 'grasp' with our minds. However, *meaning is completely determinate* in this fundamental primitive language, and this is done without the need for rigid 'rules' governing the game. Understanding the meaning of an action in these primitive cases does not depend on pinning down some

interpretation of a rule or recognizing any regularity between the current instance of the rule and our past experiences of different actions and rules.

Wittgenstein challenges us, in PI 303, “Just try – in a real case – to doubt someone else’s fear or pain”, but this is not a typical bet. Usually, bets can fall either way. Here Wittgenstein suggests that it is *impossible* to doubt someone’s pain *in a real case*. This challenge is *not* a simple expression of the ‘extreme likelihood’ that someone else is in pain in a ‘real’ case; it is rather a novel act of defiance against the traditional distinction between action and meaning. We cannot coherently ask how *this* action ‘accords’ with *this* rule because the meaningfulness of the act and the act itself are tied together so thoroughly that you cannot coherently ask about one in separation from the other. Therefore, we are given no *real* room for doubt, for doubt can only exist where a possibility of error exists. In these cases, there is no room for error because of a two-fold mutual dependence. On the one hand, the action is the sole conveyor of the meaning, while on the other hand we would have an entirely different action if the meaning altered.

So where does this leave us with the issue of normativity? The whole idea of justifying our actions in this primitive language has not been given a clear sense because of the mutual dependence between meaning and action in these cases. Actions cannot be brought into accord with a rule because they are inextricably bound together. This negative response is only one side of the story. The positive response tells us that this primitive language is one with which we are instinctually familiar. Consequently, we never make false moves in this language-game because the meaning of the gestures and sounds within it are clear to us on a deep intuitive level. Our primitive language needs no ‘justification’ because our ability to play it cannot coherently be called into question.

III. Extending Our Results

These meditations may have established a plausible account of our primitive language, but we have mastered a much more intricate extension of this language, which has presumably developed throughout our evolutionary history. Wittgenstein explicitly suggests that this kind of development is the case:

Being sure that someone is in pain, doubting whether he is, and so on, are so many natural, instinctive, kinds of behaviour towards other human beings, and our language is merely an auxiliary to, and further extension of, this relation. Our language-game is an extension of primitive behaviour. (For our *language-game* is behaviour.) (Instinct). (Z, 545)

It still remains to be seen how this ‘extension’ is possible. Wittgenstein emphasizes the need for training and regularity to establish what we call ‘language’ today. For example, Wittgenstein ends PI 207, the section about gagging a member of a community whose sounds were not regularly connected with their actions, by saying, “Are we to say that these people have a language: orders, reports, and the rest? There is not enough regularity for us to call it ‘language.’” Of course, appealing to the concept of ‘regularity’ seems problematic at first, because we largely avoided it when describing our primitive language. Wittgenstein is aware of this, and follows PI 207 with a response:

Then am I defining “order” and “rule” by means of “regularity”? – How do I explain the meaning of “regular”, “uniform”, “same” to anyone? – I shall explain these words to someone who, say, only speaks French by means of the corresponding French words. But if a person has not yet got the *concepts*, I shall teach him to use the words by means of *examples* and by *practice*. – And when I do this I do not communicate less to him than I know myself.

In the course of this teaching I shall shew him the same colours, the same lengths, the same shapes, I shall make him find them and produce them, and so on. I shall, for instance, get him to continue an ornamental pattern uniformly when told to do so. – And also to continue progressions. And so, for example, when given: to go on:

I do it, he does it after me; and I influence him by expressions of agreement, rejection, expectation, encouragement. I let him go his way, or hold him back; and so on.

Imagine witnessing such teaching. None of the words would be explained by means of itself; there would be no logical circle. (PI, 208)

The first two paragraphs are initially confusing. Wittgenstein seems to be alluding to the techniques that he criticized elsewhere in the PI, but these were criticized in the context of critiquing TA. We should not think that this kind of teaching allows the student to grasp the rule by some private mental operation of their minds – in either a Russellian or Fregean way or through developing ‘dispositions’. The last two paragraphs indicate a vital change in direction. The teacher *influences* the student *using primitive expressions*, such as rejection, agreement, expectation, and others. *These* expressions are already meaningful for the student, and they are able to use them as a solid foundation for their learning. This avoids problems about logical circularity because we have taken the meaningfulness of our primitive language as a brute fact. Now, for example, if we consider the student learning the series of even integers, her teacher gives her a number of examples, lets her try to continue, and then either expresses encouragement or rejection. These reactions are the kinds of primitive behavior that the child understands already, and so progress can be made during the child’s training.

A natural question arises at this point, because it is not as clear how these actions and the meaning the teacher is trying to convey are inextricably bound together. How can any number of these primitive ‘meaning-actions’ give rise to the kind of meaning we desire when speaking the language game of mathematics? We are trying to teach the student the *meaning* of the expression “ $2n$ ”, which is presumably different from any of the primitive gestures we have been discussing, *but how is this done?*

The answer to this question lies in a more thorough discussion about the relationship between rules and meaning. We need to make a careful distinction between *what we call “rules”* and the meaningfulness of particular actions. In our vastly

complicated language game of English, we call things like “2n”, or various other grammatical rules, “rules”, and it is precisely this aspect of our language that beguiled us into adopting TA. We have to follow, or ‘be in accordance with’, a *rule* for our actions to be meaningful – how this is established is another matter. Here the rule, there the action, and the particular relationship between them will determine the meaningfulness of the act. But this is exactly the kind of picture we must reject, on the pain of facing the devastating problems raised in the first chapter. Rather, the things we call rules are *linguistic expressions* that serve as what Wittgenstein calls ‘sign-posts’ for us. They are also the things we use to ‘justify’ our behavior, which brings us to the vital distinction between meaningful actions and rules:

“But how can a rule shew me what I have to do at *this* point? Whatever I do is, on some interpretation, in accord with the rule.” – That is not what we ought to say, but rather: any interpretation still hangs in the air along with what it interprets, and cannot give it any support. Interpretations by themselves do not determine meaning.

“Then can whatever I do be brought into accord with the rule?” – Let me ask this: what has the expression of a rule – say a sign-post – got to do with my actions? What sort of connexion is there here? – Well, perhaps this one: I have been trained to react to this sign in a particular way, and now I do so react to it.

But that is only to give a causal connexion; to tell how it has come about that we now go by the sign-post; not what this going-by-the-sign really consists in. On the contrary; I have further indicated that a person goes by a sign-post only in so far as there exists a regular use of sign-posts, a custom. (PI, 198)

The interlocutor in this passage assumes the traditional distinction between meaning and action that we have rejected. There is no such ‘act of accordance’ between an action and a rule; indeed this model loses all explanatory power when its logic is followed through:

This was our paradox: no course of action could be determined by a rule, because every course of action can be made out to accord with the rule. The answer was: if everything can be made out to accord with the rule, then it can also be made out to conflict with it. And so there would be neither accord nor conflict here.

It can be seen that there is a misunderstanding here from the mere fact that in the course of our argument we give one interpretation after another; as if each one contented us at least for a moment, until we thought of yet another standing behind it. What this shews is that there is a way of grasping a rule which is *not* an *interpretation*, but which is exhibited in what we call “obeying the rule” and “going against it” in actual cases. (PI, 201)

Much needs to be said to elucidate these infamous passages, but they are an ideal starting point for sketching a picture of how we “obey a rule” and “go against one” in particular scenarios. A more robust account of this position requires the further discussion of the role of ‘training’ and ‘custom’.

There are at least two things needed for the acquisition of language: practice through training and an existing custom within a community that serves as a background for this teaching. We will first consider the role of training, but we should contrast training with our primitive community to emphasize the differences in the scenarios. Compare the classroom setting to that of the primitive group we have been considering. In the primitive group, the actions of its members are fairly *loose*, I want to say. Each member interacts with the others in various ways: playing, fighting, eating, mating, walking, and searching, all the time interacting with whom they please. There is no emphasis placed on the repetition of certain sounds in certain scenarios because all of their interactions are, for the most part, haphazard. The group takes life as it comes to them, dealing with problems as they arise without giving attention to regulating the sounds they happen to make to each other. There is certainly no such thing as ‘training’ present in any of the groups *activities*, and, as we have seen, it is precisely *only in action* that meaning can take hold.

We witness very different activities in the classroom. We can see that there is much more going on here than mere primitive gestures would allow. While there are elements of the primitive language present in these teachings, the focus on repetition and the way old words are put together into new explanations introduce foreign situations to the students. Using these techniques of repetition and examples, the teacher *forces* a new

kind of activity upon the students. The teacher *uses* the series in front of the students, all the while offering explanations to supplement these uses. These rigid techniques are alien to the loose environment of the primitive community. However, the students are not left in the dark because the expressions of encouragement, rejection, and others that the students understand are constantly employed by the teacher to lead them into novel settings. In this way, the teacher *demonstrates* meaning *by acting*,

Every sign *by itself* seems dead. *What* gives it life? – In use it is *alive*. Is life breathed into it there? – Or is the *use* its life? (PI, 432)

“Everything is already there in ...” How does it come about that this arrow → *points*? Doesn’t it seem to carry in it something besides itself? – “No, not the dead line on paper; only the psychical thing, the meaning, can do that.” – That is both true and false. The arrow points only in the applications that a living being makes of it.

The pointing is *not* a hocus-pocus which can be performed only by the soul. (PI, 454)

Yes: meaning something is like going up to someone. (PI, 457)

It is in the application of the series that meaning gets its hold. PI 457 describes this in a brief metaphor as ‘going up’ to the students. The action is right in front of the students, and so the meaning is too. The teacher culminates her lesson by showing the students what we call ‘a rule’, which essentially serves as a sophisticated *gesture* that functions as a ‘sign-post’ for the students.

We can do more to clarify this picture. During the course of her teaching, we see that the teacher *entirely exhausts* what she knows when all of her examples and explanations run out. Wittgenstein has several quotes endorsing this idea:

What does it mean to know what a game is? Isn’t my knowledge, my concept of a game, completely expressed in the explanations that I could give? That is, in my describing examples of various kinds of game; shewing how all sorts of other games can be constructed on the analogy of these; saying that I should scarcely include this or this among games; and so on. (PI, 75)

“But do you really explain to the other person what you yourself understand? Don’t you get him to *guess* the essential thing? You give him examples, – but he has to guess their drift, to guess your intention.” – Every explanation which I can give myself I give to him too. – “He guesses what I intend” would mean: various interpretations of my

explanation come to his mind, and he lights on one of them. So in this case he could ask; and I could and should answer him. (PI, 210)

The teacher leads her students by means of initially simple examples, namely the beginning of the series, gives some kind of explanations of the form, “And now repeat the same procedure at each step”, gives examples of false moves perhaps, and so on. If she comes to a point where there is *no further activity* she can *do*, she has exhausted all of her knowledge. This reinforces Wittgenstein’s emphasis on action and helps us understand how he intends to restrict meaning to this external sphere. We need to keep in mind the fact that the teacher does not perform any private act of ostensive definition *that determines the meaning* of “2n”. Every explanation of the meaning of “2n” the teacher can give to herself she can exhibit in her actions, and so if all of these explanations are exhausted in action, she has no further ‘private’ understanding to appeal to.

Two questions should be noted before continuing: (1) what role are interpretations playing here and (2) what are the exact resemblances that this account bears to the dispositional account discussed in chapter one. Concerning (1), we indeed need to be careful about the role interpretations play in this training, lest we implicitly commit ourselves to TA. As mentioned in PI 210, there is a role interpretations *can* play in mastering language. Various interpretations can occur to the mind of the student throughout this training, *but they hang in the air* with what they interpret – they do not determine meaning. Wittgenstein is not concerned with eliminating the idea that we interpret things at all and so he wants to acknowledge the legitimacy of our natural inclinations to stress the importance of our ability to interpret when learning our language. Indeed, he still refers to them at points, but what is *essential for meaning* here has to do with *actions*, and nothing to do with anything entirely private. *Interpretations*

can be brought into accord or conflict with any action, and so they are impotent when it comes to determining meaning. Wittgenstein explicitly discusses the superfluity of interpretations in his discussion of pain, but he makes this point more subtly in other cases:

“Imagine a person whose memory could not retain *what* the word ‘pain’ meant – so that he constantly called different things by that name – but nevertheless used the word in a way fitting in with the usual symptoms and presuppositions of pain” – in short he uses it as we all do. Here I should like to say: a wheel that can be turned though nothing else moves with it, is not part of the mechanism. (PI, 271)

[...] Well, suppose that a picture does come before your mind when you hear the word “cube”, say the drawing of a cube. In what sense can this picture fit or fail to fit a use of the word “cube”? –Perhaps you say: “It’s quite simple; -- if that picture occurs to me and I point to a triangular prism for instance, and say it is a cube, then this use of the word doesn’t fit the picture.” – But doesn’t it fit? I have purposely so chosen the example that it is quite easy to imagine a *method of projection* according to which the picture does fit after all. (PI, 139)

[...] So our ‘belief that the picture forced a particular application upon us’ consisted in the fact that only the one case and no other occurred to us. “There is another solution as well” means: there is something else that I am also prepared to apply such-and-such a picture, such-and-such an analogy, and so on.

What is essential is to see that the same thing can come before our minds when we hear the word and the application still be different. Has it the *same* meaning both times? I think we shall say not. (PI, 140)

I hearken back to earlier sections in the PI to draw a connection between the cases of pain and other words in our language because I do not want to bring in additional confusions just yet about mental states. Let us restrict our questions to rules and meaning for now. Images of leaves, the color green, cubes, and even pain are all the analogous counterparts for whatever private interpretation occurs to the students learning the mathematical rule “ $2n$ ”. However, as we can see in each case, these are completely irrelevant when we are concerned with *actions*, when we are concerned with *meaning*. PI 271 is vital: the interlocutor asks a particular instance of the more general question: “What if completely different interpretations occurred to the student all the time and yet their actions never revealed their varying interpretations?” The answer is that if their actions are in line with

the use of a word, we can dismiss their private interpretations as irrelevant. This is brought out in PI 140 even more: consider two people who have the *same* interpretation occur to them, but who nevertheless use the word ‘cube’ differently. For example, perhaps both think of the ‘image’ of a cube in their minds, but then one projects this onto rectangular prisms while the other projects it onto cubes. Wittgenstein suggests that these people have the same meaning for the word ‘cube’, which seems highly plausible. This conclusion suggests that the existence of interpretations remains irrelevant for meaning and *does not* commit us to TA.

The interlocutor’s question remains weighty *only if we accept the traditional distinction between meaning and action*. If we have any residual urges to gasp at such a response, it is only because we are trying to hold onto our old biases. For example, one might want to accuse Wittgenstein of some form of verificationism. “He is simply saying that we can only know what the student means based on their actions. Since their actions are finite, there will always be a possibility that they are following a different rule, for, as he himself was so apt to point out, these can accord with any number of rules.” But this response associates directly with the assumption we deny. There is *not* a possibility for the student to ‘mean one thing’ privately and *do* something else publicly. Wittgenstein does not ascribe to either a Russellian/Fregean or verificationist theory of meaning.

With the role of interpretations settled, we can turn to the question about the relationship between this account and the dispositional theory critiqued in chapter one. This picture of training bears more than a superficial relationship to the dispositional account. However, we must first clarify that the dispositional response in chapter one was offered as a response to the rule-following problems *by someone who adopts TA*, and we

showed how their account fails to meet their own requirements. Now that we have *rejected* TA, we can reevaluate possible benefits of some form of the dispositional account, while avoiding implicit commitment to TA.

Recall the thrust of the problems with the dispositional response: the finitude of our dispositions leaves the possibility open for us to be following an infinite number of possible rules and, since *it is the rule* that normatively dictates infinitely many applications of the word, this response merely establishes that whatever *we feel* is the correct application of a word *is* correct. This prohibits the possibility of mistaken applications of a rule, which in turn prohibits the possibility for teachers to correct their students in the course of their studies. However, both the dispositional response *and* its criticisms assume TA. The dispositional response intended to show how *our actions can be brought into 'accord' with an abstract rule*, and the problems raised against it showed how it failed to do so. Since we have rejected TA, any similarities that the role of training might bear to the dispositional account should be viewed as innocent until proven guilty of the earlier problems levied against dispositionalism.

With this confusion averted, we must turn back to training, recalling to mind crucial passages of Wittgenstein's quoted earlier (namely: PI 198, 201, 208). One of the most important aspects of this training that might easily be overlooked is that the training *takes time*. A student does not learn the meaning of "2n" instantaneously. She must go through several examples over several days (or longer) under the guidance of her teacher. This process could be easy or arduous, but it will always be a kind of *development* over time. We come to master other words in our language in similar ways; indeed, the mastery of the English language takes many years (and when can we say to have *finished*

mastering English?) By being present in the classroom, the student is in a constant state of growth and change towards language mastery.

This training is analogous to physical development in illuminating ways. Both time and nourishment are needed for the development of a human. It is only after a sufficient number of years that the child can do things like run, jump, eat food, and even *speak*. The entirety of the child's development was *necessary* for her to be able to do these things. Training can be seen exactly as a kind of development. We have been trained to react in certain ways to linguistic expressions in exactly the way that our development as an organism has made us able to react to physical stimuli in certain ways. For example, we can digest particular kinds of foods, and not others, we cannot breathe under water, but can otherwise, all because of our development. Similarly, we react to "2n" in *this* way, we call *these* things "green", because of our training. Furthermore, just as our development entails novel multifarious interactions between cells, so too does our training put us in foreign environments with new interactions between known primitive gestures that move us towards language mastery.

This comparison illustrates the ways in which the meaningfulness of expressions in our extended language transcend the meanings in our primitive language. We start with relatively simple interactions among cells, but after a time we have a highly complex organism whose physical nature and capabilities go far beyond the limitations of the simple cells from which it originated. It also shows us how this account relates to dispositionalism. We have developed particular ways of reacting to sign-posts through our training over time, and so in a sense we might want to say we are "disposed" to act in certain ways towards them. However, the problems about normativity leveled against

dispositionalism in the first chapter do not arise in this account. The reasons dispositionalism failed there was because of a misguided assumption about meaning (TA) and the resulting normative requirements placed on meaningful language acquisition.

This comparison between training and development does more than just offer a new way of looking at training: it helps us answer the problem of normativity. Consider the question: “Why does the human body need *water*, specifically, to function? Why can’t it use sulfuric acid to hydrate itself instead?” Now, thanks to science, we *do* have initial answers to this question. A robust response presumably involves highly technical scientific jargon that is unnecessary for our purposes. Some short answer such as “Our bodies chemistry is such that our cells require H₂O, and not sulfuric acid, to function” will suffice. Why is this answer satisfactory? Could we not just repeat our question at the level of chemistry? Why does H₂O, and not sulfuric acid, react in these ways to these other chemicals? Now physics steps in with some explanation about how electrons and protons interact with each other, but what do we say to the interlocutor who asks, “Why does it take only two electrons to fill the first orbit of an atom, and not three?” This question seems to be a product of a deep misunderstanding. The only appropriate response to this question is simply that *this is how things are*. If you persist to ask “Why?”, I am inclined to say that you don’t understand what appropriate questions look like. We have exhausted all possible justifications, but physicists are not thereby somehow ignorant of what they are talking about, or unjustified in making the physical theories that they do.

This shows us exactly how we are to answer the problem of normativity. We have seen that normative requirements cannot be as rigid as those demanded by TA, but we

were still left with the question, “How do we know what the *right* way to follow any sign is?” Our answer is *this question itself becomes meaningless* after we have exhausted all of our justifications. This answer seems to be completely in line with what Wittgenstein had in mind:

“Let me ask this: what has the expression of a rule – say a sign-post – got to do with my actions? What sort of connexion is there here? – Well, perhaps this one: I have been trained to react to this sign in a particular way, and now I do so react to it. (PI, 198)

How can he *know* how he is to continue a pattern by himself whatever instruction you give him? – Well, how do I know? – If that means “Have I reasons?” the answer is: my reasons will soon give out. And then I shall act, without reasons. (PI, 211)

“How am I able to obey a rule?” – if this is not a question about causes, then it is about the justification for my following the rule in the way I do.

If I have exhausted the justifications I have reached bedrock, and my spade is turned. Then I am inclined to say: “This is simply what I do.” (PI, 217)

To use a word without a justification does not mean to use it without right. (PI, 289)

These quotes succinctly express our response. The bedrock of our training is analogous to the scientific statements of physics, while our extended training corresponds to the physical development we all undergo. If asked to justify the nature of our bedrock, we can only respond with a gape and an explanatorily void answer. Our reactions to the things we call “rules” express the nature of our bedrock:

“But the way we see it surely gives us everything!” – But that is not an observation about the segment of the series; or about anything that we notice in it; it gives expression to the fact that we look to the rule for instruction and *do something*, without appealing to anything else for guidance. (PI, 228)

“The line intimates to me which way I am to go” is only a paraphrase of: it is my *last* arbiter for the way I am to go. (PI, 230)

Rules, in their multifarious forms and expressions, are simply sophisticated (audible or written) *gestures* that serve as the last arbiter for the way we should respond. Our particular response depends on our training and while we can give initial explanations or justifications for our behavior (in the same way that science can give explanations up to a

certain point), any supposed ultimate justification for *why* we respond in the way we do is doomed to incoherency.

All of these explanations about training lead us to an analogous worry that we used as a criticism against the dispositional response. There we saw that if the student's dispositions governed correct uses of a word, then effective teaching becomes impossible, for we cannot actually *correct* the student's responses at any given moment. Does a similar worry arise with this account of training? These questions hinge on our understanding of Wittgenstein's use of 'custom' and his notion of community, as well as what I call his *positive* private language argument, which brings us to our next chapter.

Chapter Three:

The Private Language Argument and the Problem of Other Minds

The picture of meaning sketched in the previous chapter directly aids our understanding of Wittgenstein's infamous 'private language argument' (PLA), which will in turn buttress Wittgenstein's dissolution of the mind/body problem. This chapter is accordingly divided into two parts: the first pins down the PLA while the second applies its conclusions and our previous lessons to address the status of 'mental states' and the problem of other minds. The first section argues that Wittgenstein denies that there can be any private language *if* one thinks that "privately following a rule" means performing some private mental act of grasping or ostensive definition. However, Wittgenstein *does* accept that there can be a kind of 'private language' if by this we are simply considering a physical person in isolation. These are two, albeit interrelated, conclusions that follow from the different meanings of the word "private" and the results of the previous two chapters. I give the name "negative PLA" to the former and "positive PLA" to the latter. The discussion of the positive PLA will address the questions about custom that we left off with in the end of the second chapter. The second part of this chapter starts by arguing for a rejection of the Cartesian distinction between mind and body, a move methodologically similar to our earlier rejection of the traditional assumption about meaning (TA). The chapter closes with a new conception of mental states and their relation to human beings. With this the stage will finally be set for our discussion of solipsism and the self.

I. The "Private" Language Argument

The previous chapter left us with some startling results, the main of which claims that there is no distinct division between meaning and action. We have also seen that traditional theories of semantic meaning based on TA should be rejected on pain of suffering devastating problems advanced by the rule-following considerations. Many such views rely not just on our ability to *privately* correlate words with meanings, but that this act of mental ostensive definition *determines* the meaning of all of our words. These views cohere with many of our intuitions. It seems to be a trivial truth that I can inwardly say a word to myself and give it a meaning, even without any particular external accompaniment. Moreover, after having done so, I could presumably use these words meaningfully, even though no one else would understand me. In this broad framework, many people endorse the possibility of a ‘private language’, and by this they mean that the meaningfulness of our words depends *only* upon the ability to privately ostensively bestow words with ‘meanings’. We considered two influential accounts of these ‘meanings’: one where they were sensations and the other where they were abstract objective entities. For our purposes, the essential part of these theories lies in their commitment to this *private* mental act of meaning assignment. The sense in which these traditional theories use ‘private’ is one of the senses of the word that Wittgenstein considers in the *negative PLA*.¹⁴

¹⁴ Before spelling out Wittgenstein’s refutation of the possibility of a ‘private language’ in *this* sense, I would like to note that I am not directly concerned with *where* this argument is to be found in the PI, whether before PI 243 or after. The later Wittgenstein was a decidedly *unsystematic* philosopher, and the PI clearly reflects this style of thinking. Therefore, I will use quotations from both before and after PI 243 to explain the PLA, for both sections seem to offer important points for the PLA’s defense. I am not concerned with a detailed exegesis of the PI, and so endorsing the view that the PLA is wholly contained within one section as opposed to another is irrelevant for our purposes.

With this in mind, we can begin by noting that there are strong indications that parts of the negative PLA are contained in passages prior to PI 243, and PI 202 is a prime example:

And hence also ‘obeying a rule’ is a practice. And to *think* one is obeying a rule is not to obey a rule. Hence it is not possible to obey a rule ‘privately’: otherwise thinking one was obeying a rule would be the same thing as obeying it. (PI, 202)

In light of the first two chapters, we can see that there will be significant problems facing an advocate of this kind of private language. There is no distinction between meaning and action, and ‘interpretations’, which here can be seen as acts of private mental ostensive definition, occurring to the speaker are irrelevant for meaning. Wittgenstein puts forth the negative PLA to combat those who would draw a distinction between inward *private* ostensive definition and actions that correspond to them. Wittgenstein fights this conception of meaning on its own grounds; the model commits itself to self-defeating problems about normativity and deep relativistic problems about meaning. A detailed discussion of the negative PLA lies in the first chapter, and so the short summary is presented here by way of contrast with the positive PLA.

What remains, and what is more interesting, is the positive PLA, which will elaborate our discussion of meaning from the second chapter. Wittgenstein hints at the possibility of there being something we might call a “private language” and Baker and Hacker suggest that meaningful language does not have to have anything to do with a community of people:

[...] sounds which no one else understands but which I ‘*appear to understand*’ might be called a “private language”. (PI, 269)

[...] in its original contexts in both manuscripts the remark that following a rule is a practice has nothing directly to do with *social* practices. Its exclusive concern is with the fact that rule-following is an *activity*, [...] (Baker and Hacker 1984, p. 16)

When we considered our primitive community, we saw that their actions *were* meaningful, despite their sound's irregular correlation with their circumstances. But before going further, we need to distinguish between two important scenarios that might be called instances of "private language". Firstly, we can consider a physically isolated person *who has not been raised in a society* – i.e. who has not undergone any of the training we have received. Secondly, we could consider someone like Robinson Crusoe, a man raised in a community of language users who becomes marooned on an island.

On the authority of Baker and Hacker, Wittgenstein apparently *did* discuss the case of Robinson Crusoe explicitly:

Interestingly, Wittgenstein did explicitly discuss Robinson Crusoe in his notebooks. MS 124 has an early version of *Investigations* S 243(a), a discussion of the imaginary monologists, whose language is translatable by the explorer. Couldn't we imagine people who speak only to themselves? In that case, Wittgenstein responds, each person could have his own language. There could be men who know only language-games that one plays by oneself, viz. ordering oneself, telling oneself, asking and answering oneself, etc. How they learnt their language is here irrelevant, he adds. An explorer who observed the behaviour of such monologists could translate their languages. Later Wittgenstein remarks that the private language that he has described above is one which Robinson could have spoken to himself on his island. If anyone had observed him, he could have learnt this language. For the meanings of the words of his (contingently) private language are shown in Robinson's *behaviour*. (Baker and Hacker 1984, p. 41)

And in the footnote to this section they add:

Crusoe could certainly play language-games by himself, Wittgenstein remarks. If one secretly observed his sign-using activities, and *if one discerned in them certain kinds of complex regularities*, one would rightly judge him to be using a language of his own. (Baker and Hacker 1984, p. 41, my emphasis)

There are a number of things to discuss from these quotes. For starters, I made the distinction between Robinson Crusoe, *a man trained in a society*, from someone raised isolated from any community of people, but from these quotes it appears that Wittgenstein considered *how* the solitary person learned their language to be irrelevant. While this might be initially misleading, it does nothing to hinder my distinction. The distinction is meant to contrast someone who has been *trained*, by whatever means, with

someone who has not. It is simply easier for us to envision training happening within a community of people.

Wittgenstein gives at least one somewhat specific requirement for the word “language”: what we call a “language” is a series of words with enough regularity for us to be able to translate it into our own language. This is directly related to the fact that we said learning what we call “rules” involves a training that forces *regularity of action* on a student, and it is clear that Wittgenstein emphasizes the need for training: “Following a rule is analogous to obeying an order. *We are trained to do so*; we react to an order in a particular way.” (PI 206, my emphasis) What Wittgenstein does *not* specifically mention is that this training *has* to be done by a *person* (or by a community of people).

Wittgenstein considers it *logically possible* for someone to receive training in a language from other places, such as their environment or their interactions with other animals, and, in so far as this is possible, we can consider it possible for there to be a “private language”.¹⁵

¹⁵ Since I mentioned ‘logical possibility’, Baker and Hacker push this point further by noting Wittgenstein’s opinion on creatures born with the ability to speak a language readymade:

It is important to note that Wittgenstein countenances the logical possibility of creatures being born with the ability to speak a language. [...] *How* one has learnt or acquired a language is irrelevant to an account of *what* one has learnt. [...] Is his [Robinson Crusoe’s] continuing to be able to do so dependent on the history of his acquisition of his linguistic skills? That seems inconsistent with the principle that ‘Teaching as the hypothetical history of our subsequent actions ... drops out of our considerations’ (*Blue and Brown Books*, p. 14).

[...] Wittgenstein concludes that to describe the language of a people is to describe a regularity of their behaviour, and to describe a language which someone speaks only to himself is to describe a regularity of *his* behaviour, and not something which can happen only once. (Baker and Hacker 1984, p. 20-1)

I do not wish to digress too far into the differences between ‘logical’ and ‘physical’ impossibility for Wittgenstein. It suffices to say that I think Wittgenstein considers it *physically* impossible for someone to be born speaking a language already, although it is not logically impossible. (This is most likely related to his insistence that there is no *logical* impossibility for us to feel another person’s pain (i.e. PI 302).) Training is *physically* necessary *for us* to learn language, but in what sense does it “drop out” of our considerations as irrelevant? It drops out *only in so far as* the observation of someone born

There are marked differences between this kind of private language and the one attacked in the negative PLA. First of all, the language of the trained solitary individual is meaningful, just as marooned Robinson Crusoe's language is meaningful. *Our language games are behaviours* – they are actions, and in so far as they are nuanced and highly developed actions is the extant to which they are meaningful. What we call “language” is only possible with a certain consistency in actions, which is of course *possible* for a solitary individual to develop. This is directly opposed to the view of private language attacked in the negative PLA, where the meaningfulness of an individual's words lied in an act of private mental ostensive definition. Robinson Crusoe *can* talk to himself, give himself orders, and so on, because he has already mastered the technique of English; he has already been trained to do these sorts of things. However, this language is not in principle inaccessible by anyone else.

This settles how Wittgenstein intends us to understand ‘custom’ or ‘practice’. We *can* think of a custom that supports the training of a language to consist in a community *of people*. In fact, this is undoubtedly the most common way people do learn language. The environment in which most of us acquire language consists of extended periods of interaction with our parents, as well as other people, and these factors typically serve as the regulating force needed for training. The inherited trainings your parents and larger

speaking a language does not require an analysis of their previous linguistic history (indeed, because it is non-existent) for us to observe that they are following rules – i.e. that their actions are following certain regularities. We *could* translate this infant's behavior; it *would* be a language. But think back to the analogy of development: this ‘logically possible’ case is analogous to someone being born immediately as an adolescent. I *suppose* that is ‘logically possible’ (and if it is not, then we have nothing to worry about). But now consider the role of explanations: in normal cases of language development, we can legitimately ask: “Why do you respond in the way you do upon the order ‘x’?” The answer to this appeals to our previous linguistic history, which of course *must end* at some point – when we have no further explanations. In this strange logically possible case, *we hit bedrock immediately* when we ask the infant “Why do you do what you do?” Of course, the lack of explanations does not force us into saying the infant's actions are without meaning – it is just that in this case we have no explanations to offer.

community received provide a kind of ‘custom’ in which you learn language. The concept of ‘custom’ here is deeply analogous, and probably overlaps with ‘culture’ and the extent to which cultures vary gives us a good picture of how variable the customs in which we learn language can be. However, the question naturally arises: do we all have our own *different* ‘cultures’? In what sense can we be said to mean the same thing by our words? Moreover, if I speak English and hear someone misuse a word, what justifies me in correcting her?

The answer to these questions requires us to backtrack through our discussion in chapter two. Wittgenstein is sensitive to these problems, which he ultimately answers by an appeal to primitive language and ‘forms of life’.

Following a rule is analogous to obeying an order. We are trained to do so; we react to an order in a particular way. But what if one person reacts in one way and another in another to the order and the training? Which one is right?

Suppose you came as an explorer into an unknown country with a language quite strange to you. In what circumstances would you say that the people there gave orders, understood them, obeyed them, rebelled against them, and so on?

The common behaviour of mankind is the system of reference by means of which we interpret an unknown language. (PI, 206)

“So you are saying that human agreement decides what is true and what is false?” – It is what human beings *say* that is true and false; and they agree in the *language* they use. That is not agreement in opinions but in form of life. (PI, 241)

The foundation of all our complex language games rests on primitive language, and if we came upon someone who reacted differently to the gestures and sounds that constitute our primitive language, we can only say that they have a dramatically different form of life. This is similar to meeting someone who reacts to the pointing of a finger by looking in the opposite direction (PI 185): if you think that the ‘point’ of the pointing is to direct your attention in the opposite direction, then it seems that we must throw up our hands and resign ourselves to silence. This answer does not indicate a point of defeat; it is supported by deep natural inclinations and tendencies that can be brought out by means

of an example. Consider trying to teach a child to look in the opposite direction when given a gesture of pointing with the following caveat: you *cannot* use English (or another language) to do so. It seems highly implausible that you could get the child to look in the opposite direction under such constraints. This simple consideration shows why we should feel safe in adopting such a premise as a ‘brute fact’.

Returning to our more complex language games, we can consider another thought experiment to demonstrate our answer. Suppose I misuse the word “green” when describing the color of my hair to my friend (and that *would* currently be an inappropriate use), who proceeds to gape at my apparently sincere statement. After her shock wears off, she proceeds to try to convince me to retract my statement, presumably by means reminiscent of the training she underwent concerning the word “green”. Perhaps some of these initial explanations fail because they rest too much upon our extended language games, but ultimately she will do something that belongs to our primitive language. As a *human being*, it is impossible for me to misunderstand her at this point. Someone (or more properly something) who did not understand these gestures would not share the same form of life. (“If a lion could talk, we could not understand him.” (PI, p. 223)) The conclusion of this thought experiment claims that we *can* legitimately correct people using our language by an ultimate appeal to our primitive language. The common behaviors of humanity are the means by which we found any language, which explains why they are our only means of translating languages. Moreover, appealing to primitive language gives us a means of explaining our meaning of the word “green” to someone who misuses it. Three things are possible after we give such an explanation. The person could retract their former use of the word, followed by a correct use of the word. The

person could respond, “Oh that is not what I mean by ‘green’. I guess I use ‘green’ to mean what you call ‘brown’”, in which case we simply use different words for the same meaning. Here we are able to understand each other, although a process of translation is necessary to do so. The third possibility is that the person lacks the ability to learn the meaning of the word, which is certainly possible. Again, whether or not the person correctly learns the meaning of the word ‘green’ will manifest itself and be exhausted by the person’s actions.

This example demonstrates the role of custom again. We live in a community of language speakers whose uses of words are relatively uniform, and therefore whose meanings are relatively similar. People within our community occasionally misuse words, especially when it comes to esoteric or rare words and others are able to correct one another in so far as they can all appeal to similar trainings they have received. If these appeals fail, we can teach the meanings of our words by relying on our primitive language. But we are certainly not able correct the uses of words in a different language. Moreover, there is no one set or permanent “correct” use of a word in these extended languages, for each of them merely represents a highly complicated form of life. As humans, there might be extensive overlap in the forms of life between people who speak different languages. This comes from the fact that we all share our primitive language. But the meanings of our extended language games that transcend the meanings present in our common behaviors are not constrained by any objectively normative requirements. People can only correct one another’s use of a word in so far as they all participate in the same form of life and there is no objective standard establishing correct or incorrect uses of words outside of this.

We return to our distinction between the trained and untrained solitary person. What of the actions of the solitary individual who has not been trained? Perhaps the easiest answer is “there is not enough regularity for us to call it ‘language’.” (PI 207) There is nothing preventing this individual from developing a “private language”, but it is essential to note that the absence of a language for this individual, private or not, *does not mean* that all her actions are meaningless. She still has recourse to her primitive expressions, to that which requires no training. It would only be in some strange hypothetical scenario where she was somehow ‘frozen’, unable to *act*, where she would be incapable of expressing any meaning. Our ability to act serves as the foundation for all of our language and, namely, our sensation language, which we should turn to now.

II. Descartes’ Error and the New Status of “Mental States”

The entirety of this essay has so far been concerned solely with the philosophy of language, and, more specifically, the meaningfulness of our language. This section is the first proper application of our conclusions to the realm of philosophy of mind. Naturally, there are deep affinities between Wittgenstein’s philosophy of language and his philosophy of mind, and so this section begins by comparing these two aspects of his thought, namely by drawing the connections between his rejection of the Cartesian distinction between mind and body and his dismissal of the traditional assumption about meaning (TA), which says that there are two separate and discrete entities necessary for meaningful discourse: physical linguistic expressions and the ‘meanings’ that attach to them. This is followed by a quick summary of the traditional problem of other minds, and

the chapter ends with a description of Wittgenstein's understanding of "mental states" and their relationship to our notion of a human being.

This essay began with a broad, somewhat platitudinous statement about the interpenetration between all fields in philosophy, but here we have clear evidence for this truth. The Cartesian distinction between mind and body has dominated the philosophy of mind since the seventeenth century. Let us call this "the Cartesian Assumption" (CA), which more precisely claims that there are *two separate and discrete* entities, the 'mind' and the 'body', that combine together in some way to form a human. It is philosophy's business to establish how this is done.¹⁶ There are clear structural similarities between the theories of those philosophers of mind who adopt CA and philosophers of language who adopt TA. One says, "Here the mind, there the body" while the other says, "Here the meaning, there the (physical) use of the word", and both then set to work figuring out the relationship between the elements of each dichotomy. TA divorces meaning from action, thereby leaving action as something impotent without its corresponding meaning. In the same way, CA divorces mind from body. Traditionally, not all bodies have minds, but certain ones, namely human bodies, may possess them. Adopting CA causes a division between bodies, which become mere physical, senseless objects, and minds that are 'in' them somehow. But the relationship between these philosophies goes beyond structure. The CA gives us a distinctly *private* picture of the

¹⁶ To say the least, we would go too far a field to even briefly discuss the multiple influential theories based on this assumption. Moreover, in the spirit of fairness, we should note that this distinction can lead to *some* solutions to philosophical problems, depending on the brand of Cartesianism you adopt. For example, questions of personal identity gain some headway with this distinction. How are we the same people over time even though our bodies are in constant flux? Our *minds* stay the same while our bodies change, and having the same mind is what is necessary for personal identity. Needless to say, there are numerous problems that riddle such a response, as there are for many Cartesian theories. In any case, while these and related problems are undoubtedly interesting, I will restrict our discussion to the traditional problem of other minds that faces anyone adopting CA.

mind, and it is in this realm that private ostensive definitions supposedly take place. Philosophers adopting both CA and TA point to the mind as the bearer of our sensations and the source of our understanding. Appealing to the privacy of our minds explained things like why we felt there could be private languages, how we could grasp the rule “2n”, how we knew the meaning of the word “red” and various other problems.

Before jumping to our rejection of CA, we should summarize the traditional problem of other minds to gain a clear view of our target. Here is the problem: I have private access to my mind, a kind of access to which no one else is privileged. I know what “red” means based off of my own private sample of the color in my mind; I know what “pain” means based off of my own experience of pain. After various experiences, these meanings are definitively *pinned down by private ostensive definition*, the very process we have shown to be incoherent. Adopting such a method leads to the following problems. If meanings are established privately in each individual’s minds, how do I know that another person means the same thing I do by the word “pain”? And what about “red”? And what about the rest of language? In fact, these worries extend beyond the meanings of our words: since I cannot directly perceive the existence of someone else’s mind, how do I know that they have minds at all? I merely see their bodies, which are separate from their minds, and so how can I be sure that they are not sophisticated robots? Am I the only *real* human being, whose mind I know to a certainty to be in its natural communion with *this* body?

The problem of other minds gains support from another important consideration. Crucial evidence for the plausibility of CA lies in the fact that there are cases of *illusion* where we might be deceived into thinking someone is in pain when they are not.

Moreover, there are cases, such as instances of acting, where we *know* the actor is not in pain, even though her behavior mimics that of a person genuinely in pain. These cases support a natural inclination to divorce the workings of the mind from physical bodies because in both we observe physical behaviors typically expressive of genuine pain in tandem with the subject not actually being in pain. CA derives strength from these inclinations: we want to say that the particular physical behaviors of a person are not enough for us to establish that we know they are in pain.

There are two things that we should consider at this point. Since we have established that our words are meaningful only in so far as they are used in certain ways coordinate with actions and circumstances, we should clarify how sensation language fits this mold. This first involves dissolving our conviction that sensations are ‘objects’ to which we have private access. The next step deals with the cases of illusion by explaining what role these cases play in our new understanding of sensations.

The CA gives us a model of sensations that Wittgenstein would call “object and designation”. Our sensation words are *nouns* after all, and so it seems proper that sensations should be objects within our minds. Now, CA implies that the meanings of our sensation words are established by private ostensive definition, but we have seen that this is impossible. The meaning of the word “pain” *does not consist in the private correlation of the word with an inner sensation*, but still an advocate of CA might insist that surely there is *something* that goes on inside us that our neighbors do not have access to. We can appeal to Wittgenstein’s famous ‘beetle-in-a-box’ passage to stamp out this temptation:

Now someone tells me that *he* knows what pain is only from his own case! – Suppose everyone had a box with something in it: we call it a “beetle”. No one can look into anyone else’s box, and everyone says he knows what a beetle is only by looking at *his* beetle. – Here it would be quite possible for everyone to have something different in his box. One might even imagine such a thing constantly changing. – But suppose the

word “beetle” had a use in these people’s language? – If so it would not be used as the name of a thing. The thing in the box has no place in the language-game at all; not even as a *something* for the box might even be empty. – No, one can ‘divide through’ by the thing in the box; it cancels out, whatever it is.

That is to say: if we construe the grammar of the expression of sensation on the model of ‘object and designation’ the object drops out of consideration as irrelevant. (PI, 293)

Since the meanings of our words lie in their use in our activities, any possible private thing that individuals associate sensations words with ‘drops out of our considerations as irrelevant’, for our sensation words are not concerned with *things*. Sensations are not *objects* of private experience; we must reject this model of ‘object and designation’ for our sensation language. Indeed, what keeps us attached to the view that sensations are objects when the meaningfulness of our sensation words would remain unaltered if there were *no* private sensation objects at all?

This is one of the primary applications of our lessons in the philosophy of language onto philosophy of mind. The essential point is that we are not committed to talking about sensations as being ‘objects’ in the ‘mind’ of a subject when we engage in the problems of the philosophy of mind. The meaningfulness of our sensation language does not consist in referring to objects of any kind and so if we want to have a *meaningful* discussion about sensations we can avoid all of the problems posed by the Cartesian assumption.

This move makes it seem like Wittgenstein comes dangerously close to behaviorism. Is he saying that sensations merely *are* physical actions, physical *behaviors*? The difference between Wittgenstein’s stance and behaviorism is subtle, but ultimately dramatically distances Wittgenstein from behaviorism:

“But you will surely admit that there is a difference between pain-behavior accompanied by pain and pain-behavior without any pain?” – Admit it? What greater difference could there be? – “And yet you again and again reach the conclusion that the sensation itself is a *nothing*.” – Not at all. It is not a *something*, but not a *nothing* either! The conclusion was only that a nothing would serve just as well as a something about

which nothing could be said. We have only rejected the grammar which tries to force itself on us here.

The paradox disappears only if we make a radical break with the idea that language always functions in one way, always serves the same purpose: to convey thoughts – which may be about houses, pains, good and evil, or anything else you please. (PI, 304)

The problem with behaviorism is that it assumes CA just as whole-heartedly as any Cartesian. The essential point of behaviorism is it reduces mental states to the same kind of *merely physical* bodily movements. These are the same bodily actions that dualists try to connect with mental states. But this view adopts a seriously misguided conception of what might be called the ‘logic of mental states’ (“Grammar tells us what kind of object anything is.” (PI 373)). Wittgenstein suggests that mental states cannot be reduced to ‘merely physical’ actions. Behaviorism advocates such a reduction, although not an ‘elimination’, of the mental.

“Are you not really a behaviourist in disguise? Aren’t you at bottom really saying that everything except human behaviour is a fiction?” – If I do speak of a fiction, then it is of a *grammatical* fiction. (PI 307)

How does the philosophical problem about mental processes and states and about behaviourism arise? – *The first step is the one that altogether escapes notice.* We talk of processes and states and leave their nature undecided. Sometime perhaps we shall know more about them – we think. But that is just what commits us to a particular way of looking at the matter. For we have a definite concept of what it means to learn to know a process better. (The decisive movement in the conjuring trick has been made, and it was the very one that we thought quite innocent.) – And now the analogy which was to make us understand our thoughts falls to pieces. So we have to deny the yet uncomprehended process in the yet unexplored medium. And now it looks as if we had denied mental processes. And naturally we don’t want to deny them. (PI, 308, my emphasis)

The “grammatical” fiction comes from our ‘first step that altogether escapes our notice’.

We adopted CA, and in doing so became committed to dividing up human action into senseless bodily movements accompanied by mental states that are ‘objects in a mind’.

Since we have denied the role any ‘private’ sensations might have in determining meaning, it looks like we should adopt behaviorism, but that is *not* what we want to do.

Rather, we should reject CA, and try to give a more accurate account of the true grammar of our sensation language.

So what *are* we left with after we reject CA? Drawing upon our earlier results, recall our primitive society. There, the meanings of primitive gestures and expressions were completely bound up with the surrounding circumstances and actions. The actions of our fellow humans are the primitive, and, in an important sense, *primary* vehicles for meaning. By rejecting CA, we are left with a merger of mind and body back into their natural home. John McDowell expresses this contrast succinctly:

One way of approaching Wittgenstein's response is to remark that such a picture [one based on CA] is attainable only by displacing the concept of a *human being* from its focal position in an account of our experience of our fellows, and replacing it with a philosophically generated concept of a *human body*. [...] In these terms, Wittgenstein's response to the sceptic is to restore the concept of a human being to its proper place, not as something laboriously reconstituted, out of the fragments to which the sceptic reduces it, by a subtle epistemological and metaphysical construction, but as a seamless whole of whose unity we ought not to have allowed ourselves to lose sight in the first place. (McDowell 1982, p. 469-70)

The sceptic McDowell refers to is someone who claims it is possible that all other humans do not really have mental states, but this doubt rests firmly upon CA. What we have left over is not a mere human *body* but a human *being*. The *primitive* actions of these beings give us cases with perfect access to their mental states because what we might call their "behaviors" are conceptually tied up with what we might call their "mental states". There is textual evidence for this stance in Wittgenstein as well:

"But doesn't what you want to say come to this: that there is no pain, for example, without *pain-behaviour*?" – It comes to this: only of a living human being and what resembles (behaves like) a living human being can one say: it has sensation; it sees; is blind; hears; is deaf; is conscious or unconscious. (PI, 281)

More needs to be said about the grammar of our sensation language. There is an important similarity between the cases of mathematical language and sensation language that will flesh out this point. We cannot reduce mathematical objects to physical objects

for the same reason that we cannot reduce mental states to mere physical behavior. All attempts to do so misconstrue the *meaning* of mathematical statements by trying to force our understanding of them into the model of ‘object and designation’, which is paradigmatic of the natural sciences but not necessarily of other language games. We naturally pursue such projects because of the superficial resemblance between the grammar of mathematical objects and of our physical objects. This is related to the distinction Wittgenstein makes between ‘surface’ and ‘depth’ grammar (PI 664): we are led by the surface grammar of our mathematical statements to think of numbers as objects in the same way that tables and chairs are objects. But this misleads us, and the differences between the two kinds of statements can be brought out quickly. We say, “She owns a chair” but not “She owns a three” or “She owns the integers”. In fact, numbers often function as adjectives, not as nouns. “She owns three chairs” is a legitimate sentence. The ontological status of numbers has been so troublesome for many philosophers because they assume that ‘only the physical’ is ‘real’, and so they proceed to try to ‘reduce’ many of our language games into statements about purely physical objects. But *numbers are not things at all*, and this does not need to worry anyone if they can accept the varying grammatical structures of our language. In the same way, we can comfortably say that mental states are not things. To do otherwise would incorrectly apply the grammar of one language game onto another, only resulting in extensive confusion.

There is another important conclusion we reach after considering the ‘logic’ of our sensation language. Despite all that has been said, one might want to insist, “Still, my sensations must be private! You do not feel what I feel when I feel pain!” This statement

emphasizes a *grammatical* point about our sensation language. “The proposition ‘Sensations are private’ is comparable to ‘One plays patience by oneself’.” (PI 248) Solitaire (also known as ‘patience’) is a game that one plays by oneself, but that reason does not mean that no one can learn solitaire from observing someone play the game. Moreover, people who know how to play solitaire understand the actions of another solitaire player. The mere fact that ‘sensations are private’ does not mean that no one knows what you are talking about when you say “I am in pain”. The language game of mental states has a meaning that is publicly visible to everyone, and the mere grammatical facts governing it do not commit us to a metaphysical picture of mental states that precludes the possibility of others meaning the same things by their words.

We acquire a better picture of the nature of our sensation language by considering cases of illusion again. We have developed a language that has phrases like “The chair is made of wood” and others like “The chair appears to me to be made of wood”. These say different things. The first says that the chair actually is made of a particular substance, while the second says that it merely appears to me that it does. We have equivalent statements in our sensation language. We can say both that “She is in pain” and “She appears to be in pain”. Let us focus on the statements about the chair. How was it established in the first place that there could be illusions about there being a chair in front of me? We have developed a language to deal with specific cases of mirages, but would it be possible for there to be *only* illusions of chairs? Perhaps one might think of this scenario as one where we just fell through chairs, couldn’t put our hands on them, and so on. In *this* sense, perpetual “illusion” would be possible, but “illusion” is in scare quotes here for a reason. Illusions are only established to be *deceptive* by reference to another

experience that establishes contrary conclusions. In other words, we determine a case to be illusory by an appeal to another experience that reveals the deceptive nature of the case in question. This statement or experience cannot be illusory itself. For example, if there were ‘nothing but illusions’ in the sense described above, there would be nothing truly illusory about chairs. Chairs would simply be objects that we could see but not touch or sit on. If it happened occasionally that people successfully sat on these kinds of chairs, *that* would be illusory.

These conclusions have immediate upshots for the case of pain. It is true that we have cases of deception or acting where we might be tricked into believing someone is in pain when they are not. However, it would be impossible for there to be *nothing but “false” moves* in our language game of sensations. Wittgenstein explicitly endorses this:

“If it is possible for someone to make a false move in some game, then it might be possible for everybody to make nothing but false moves in every game.” Thus we are under a temptation to misunderstand the logic of our expression here, to give an incorrect account of the use of our words.

Orders are sometimes not obeyed. But what would it be like if no orders were *ever* obeyed? The concept ‘order’ would have lost its purpose. (PI 345)

Correct moves in our sensation language are established by the primitive notion we have of a human *being*. The actions of these beings provide the foundation for our sensation language. These kinds of actions show us when someone is really in pain. Undoubtedly, we have a highly complex extended language game that accounts for cases of deception, but there need to be cases where someone is actually in pain (for example, the thought experiment considered in chapter two where a person falls to the floor writhing in pain) to retain our notion of the game at all. If no one’s actions actually expressed their pain, we would lose the concept of pain altogether. This would not mean simply that we might always be in error when we claim someone is in pain, but that we would not actually have the concept to begin with.

For clarity's sake, we may introduce a distinction in this argument between what I will call the metaphysical and the epistemological questions concerning pain. Metaphysical questions about pain are concerned with whether or not someone is actually in pain, regardless of our means of acquiring knowledge of this fact. Epistemological questions about pain are concerned with how we come to know someone else is in pain. In this way, there is a dependence of epistemological question upon metaphysical ones: we cannot know that someone is in pain unless it is actually possible for people to be in pain. This distinction informs an objection against the argument above. It seems that I have established the metaphysical necessity of pain – there need to be cases where people are actually in pain in order for our language to make sense – but this falls short of important epistemological requirements. While we may be in a position to potentially acquire knowledge of another's pain thanks to our metaphysical considerations, questions about how exactly we come to *know* someone is in pain remain.

This distinction brings out an important relationship between our primitive and extended languages that has been hinted at frequently. Our primitive language serves as the foundation of our extended languages and within this primitive mode of communication it seems that in certain scenarios we cannot possibly doubt the pain of others (i.e. person writhing on the floor example). But the question is: does this primitive gesture express "*pain*"? "Pain" is a word in our *extended* language game, and so we have a right to wonder whether the certainty we have in our primitive language carries over into our extended language. This point informs the objection above: our response to the problem of other minds may have established that we need cases of genuine pain for our

language to make sense, but it remains uncertain whether or not the rules governing our extended language retain the certainty present in our primitive language.

The answer to this question hinges on a dilemma: either our extended sensation language retains cases where we have no doubt as to whether someone is in a given mental state *or* we have lost access to this epistemological certainty somewhere in the process of extension. I will not definitely come down on either of these sides, but simply remark that neither present difficulties for my position. Needless to say, the first case is more desirable and might be plausible for a few reasons. When we learn sensation language, we are taught to replace exclamations of pain, such as screaming and groaning, with the words “I am in pain”. While the extent to which these expressions are exactly equivalent may be open for debate, it seems plausible that the transition from the former expression to the latter could easily retain the epistemological certainty of the former. This explains why the smile of an unweaned infant is not pretense (PI, 249): we are actually certain of the *happiness* of the child, which is a *mental state*. This brings out a *deep* relationship between our primitive gestures and sensations language. In cases of our primitive language, we may say that we know others to be in mental states because no degree of certainty was lost in extension. Indeed, the bulk of this essay supports this reading and defense of Wittgenstein.

The second alternative is more problematic, which claims that our concept of pain does not retain the epistemological certainty we have of primitive gestures. A thorough discussion of Wittgenstein’s epistemology is beyond the scope of this essay, but, more importantly, is somewhat tangential for our purposes. We should note that this alternative does not affect our metaphysical considerations, which are significant conclusions in their

own right. *There must be* genuine cases of pain, although we do not necessarily *know* which are genuine in any particular case. This solves the aspect of the problem of other minds important to our inquiry: other people *must* have mental states, on pain of undermining the meaningfulness of the language that states the problem. That being said, the remaining epistemological questions are surely important. I believe the solution to these doubts lies in a kind of transcendental argument against this epistemological scepticism, which comes from Wittgenstein's remarks in *On Certainty* and relies on Wittgenstein's notion of "criteria".¹⁷ The argument claims something like the following. The sceptical worries advanced by these epistemic questions would undermine the meaningfulness of the sceptic's thesis itself. Spatial limits prohibit a further explanation of this stance, but since the success or failure of this argument has no effect on our metaphysical considerations, we can proceed safely.¹⁸

This chapter has made several important conclusions. Building off our conclusions from chapter two, we have discussed the real sense in which people may have private languages, i.e. in a way that does not preclude others from learning them.

¹⁷ I have largely avoided discussion of criteria because of numerous controversies surrounding Wittgenstein's use of the word. I take Wittgenstein's use of "criteria" to be intentionally vague *for the purpose of dissuading* systematic analysis. He considers it an impossible task to delimit the rules governing our languages, mainly because we constantly invent new languages with new rules. There is no theoretical limitation constraining the rules we might invent, and so an attempt to pin down the exact criteria necessary for a given language seems impossible.

¹⁸ The literature on Wittgenstein's use of 'criteria' is vast to say the least. Kripke only loosely discusses it in his account (see Kripke 1982, p. 98 – 103). One influential account of this comes from Baker and Hacker's defeasibility notion of criteria (Baker and Hacker 1984). McDowell attacks this and offers 'disjunctivism' in its place. McDowell defines disjunctivism thusly: "But suppose we say [...] that an appearance that such-and-such is the case can be *either* a mere appearance *or* the fact that such-and-such is the case making itself perceptually manifest so someone." (McDowell 1982, p. 472) This stance holds that on any given case of perception we may either be experiencing an illusion (a mere appearance) or a genuine fact that such-and-such is the case. There are affinities between my view and McDowell's, although the way in which we arrive at our conclusions is different. McDowell points out significant problems with CA and then offers disjunctivism as a tenable alternative, whereas my argument suggests that some kind of disjunctivism *must* be true, not merely that it is an unproblematic option. McDowell's solution to the epistemic concerns hinted at here claims that knowledge does not require certainty. This seems to fall short of what we really want from knowledge, but I cannot comment on this further here.

We have also rejected the Cartesian distinction between mind and body, offering a new account of sensation language in its place. This account ensures that others must have mental states, although the method by which we know about people's mental states remains debatable. In any case, mental states are decidedly *not* restricted to a private realm in principle inaccessible to others. Rejecting this notion of privacy leads us to consider the self, which has often been considered the bearer of mental states. More importantly, the traditional conception of the self brings with it a notion of private communion that obtains between each of us and our respective selves. This will be attacked in both chapter four and the epilogue.

Chapter Four:

The Dissolution of Solipsism

This essay has been tracing Wittgenstein's ideas on various topics, displaying a chain of thought linking them together, and now at last we reach the culminating point in Wittgenstein's discussion of solipsism. The progression of the arguments we have been considering have an important linear structure. We began with 'rules' and adopted a traditional assumption about meaning, TA, which claims that there are two separate and discrete entities associated with our language: the physical use of the linguistic symbol and the 'meaning' that relates to it. We saw how this led to irrevocable incoherencies, and so to retain the meaningfulness of our language, we concluded that the meaningfulness of our language consists in the particular actions and circumstances of its use, with or without a community of people providing training for individuals. There are no such things as abstract 'meanings' divorced from particular circumstances and actions. Wittgenstein's discussion of meaning eliminates the possibility that any kind of private realm assists in ascribing meaning to our language. This conclusion directly implies that the meanings of the words that we consider a part of our 'sensation language' are only meaningful in so far as they are coordinate with certain actions. The meaning of these actions does not rely on the model of 'object and designation'; instead, these meanings arise not from the philosophically constructed notion of a human body but from the primitive notion of a human being, whose actions manifest 'mental states' in their fundamentally familiar way. While we did not establish conclusively our epistemic relation to cases of mental states, we determined that others must possess them in order for our language to be meaningful at all.

The layout of the chapter relies on these conclusions and follows a familiar pattern. The chapter begins by laying out the distinction between realist and solipsistic theories of the self. Two main versions of solipsism are articulated: epistemological and metaphysical solipsism. These versions are then elaborated upon by giving initial reasons why they appear appealing because of the surface grammar of “to perceive” and “I”. This is followed by the attack on solipsism, which consists in two main parts. The first claims that the ‘model of object and designation’ for the self will not work, relying on the conclusions of the previous chapters. The second part responds to the initial reasons that seemed to support the solipsist’s position. The chapter ends by summarizing our conclusions and their relation to each other.

The Traditional Assumption about the Self: Its Advocates:

The debate between solipsism and realism concerning the self should be seen as the last desperate attempt of traditionally minded philosophers to insist on the primacy of a private mental realm. These two main contenders share a fundamental assumption: both believe that the self is a kind of “object” with which we are in intimate private communion. Call this the traditional assumption about the self (SA). However, realism and solipsism differ dramatically, so we should be clear on these points of divergence before continuing.

A realist claims that each individual has a ‘self’, which has a peculiar nature (immaterial or material), that each of us refers to when we use “I” in “I am in pain”. The ontological status of these selves needs explaining, but we refer to them in the same way that we refer to physical objects or the referents of proper names. I am not directly

concerned with elaborating the realist view, mainly because the arguments levied against solipsism attack realism as well.

The solipsist denies that each individual has a self, insisting that there is only one Self, which serves as a kind of transcendental viewpoint.¹⁹ There are a few different brands of solipsism, but all of them are difficult to define precisely. Wittgenstein made several attempts in the Blue Book that seem to be acceptable definitions:

Sometimes the most satisfying expression of our solipsism seems to be this:
“When anything is seen (really *seen*), it is always I who see it”.

[...] What tempted me to say “it is always I who see when anything is seen” I could also have yielded to by saying: “whenever anything is seen, it is *this* which is seen”, accompanying the word “this” by a gesture embracing my visual field (but not meaning by “this” the particular objects which I happen to see at the moment). (BB, p. 61, 64)

Why is anyone tempted to hold such a bizarre and counter-intuitive theory? We fall into solipsism if we combine a broadly Cartesian epistemology with an emphasis on some peculiar aspects of the surface grammar of our words “I” and “to perceive”. The essence of a Cartesian epistemology holds that we have infallible access to the way things appear to us. We have perfect knowledge of our perceptions. The way things *actually* are is up for doubt, but we certainly know how things seem to us. This starting point serves as a kind of methodological solipsism: if we can gain knowledge of the external world and other minds, we have to start with what is indubitably certain to us, a la Descartes’ method of doubt.

¹⁹ The doctrine of solipsism is best understood through the mouth of a particular speaker due to the peculiar fact that, if properly understood, you should hold that the speaker falsely says “When anything is seen, it is always I who see it” because in fact *you* hold this unique position. This is a perplexing result, but not logically inconsistent. Any individual can advocate solipsism if they wish, it is just that if they meet another who believes in solipsism they will have to claim that their position is false. After this peculiar aspect of the doctrine is noted, it will be easier to refer to ‘the solipsist’ instead of the ‘theory of solipsism’ because it would be strange to say that a group of people endorsed solipsism. All of them hold the ‘same’ view, but every single one of them believes that all the others are completely misguided! In any case, I will refer to the individual ‘solipsist’ from now on rather than the ‘theory of solipsism’ to avoid this confusion.

From here, the solipsist can take two routes: epistemological solipsism and metaphysical solipsism. An epistemological solipsist comes to her position by a gross extension of the problem of other minds. We have room to doubt the existence of the entire external world, not just other minds. If the true correspondence between our perceptions and reality can be called into doubt on any given occasion, which happens in cases of illusion, then we are at risk of never having perceptions that genuinely access the world. These considerations lead the solipsist to limit her knowledge to her own experiences.

The metaphysical solipsist thinks that the epistemological solipsist moves in the right direction, but stops short of the real truth. This Self is the *source of reality* for the metaphysical solipsist and if the Self ceases to exist, so does the world. Support for this stronger claim comes from the solipsist's insistence on the peculiar reference of "I" in sentences like "I am in pain" or "I have this perception". Who has these perceptions? Or what is in pain? Since "I" in these sentences functions in a manner superficially similar to pronouns that do refer, such as "He is eight feet tall", "He is in pain", we are led to believe that there must be *something* to which "I" refers. This peculiar entity bears all the perceptions discussed above, thereby serving as the bearer of the world, in a sense. The metaphysical solipsist uses the referent of "I" to go beyond the epistemological solipsist by maintaining that the 'visual field' is all that ever exists. Whenever anything is seen, it is THIS which is seen. From the standpoint of an individual, the sort of gesture accompanying this statement is supposed to highlight more than the epistemological solipsist's claim that we have infallible knowledge of our perceptions. It is supposed to point to the 'fact' that our experiences are the only possible experiences, because, in a

sense, they are the only things that *happen*. The metaphysical solipsist says, “Every *thing* is really *just* a perception, and *I* hold these perceptions.” In this way, we are supposed to be convinced that nothing outside of our own perceptions exists.

“But what do you mean?”

With our previous chapters in mind, much should appear wrong with either of these solipsistic positions. We can apply our conclusions about mental states and cases of illusion in chapter three to the case of the epistemological solipsist. Regarding our metaphysical considerations, *there must be* cases of genuine perception of external objects. While we may not have established the exact way in which we differentiate between veridical and illusory cases, we must have some genuine cases of perception. This undermines the epistemological solipsist’s claim that we might never have genuine perceptions of external objects. There are two questions that follow from these conclusions that need answering: (1) how exactly does this conclusion establish the existence of things outside our minds and (2) what aspect of the grammar of “to perceive” tricked us into believing such a wild hypothesis?

Regarding (1), the epistemological solipsist’s conclusion was that we had to doubt our knowledge of external objects based on the fact that our perceptions can deceive us.

In response, Wittgenstein says:

The point here is not that our sense-impressions can lie, but that we understand their language. (PI, 355)

Surely our sensations can deceive us and it may prove difficult to prove to a certainty the veridical or illusory status of any given perception. However, it could not be the case that all of our sense-impressions are really illusions. Our language of perceptions accounts for

cases of deception, but this is merely one complicated rule governing our language game of sensations. It would be an improper logical inference to conclude that all of our perceptions might be illusory based only on this *grammatical* fact. The grammatical fact only allows for the possibility of *some* of our perceptions to be illusions. Implicit in this grammatical fact lies our ability to have genuine cases of the perception of external objects because it is on this more primary and fundamental aspect of our language game of sensations that the possibility of illusions rests. The upshot of this is straightforward: some of our perceptions must be veridical, so we *do* perceive things *external* to us, although we lack a detailed knowledge of which perceptions are truthful.

Part of the reason we were tempted into this kind of doubt lies in the fact that the grammar of “to perceive” allows us to speak of cases of deception but *also* because it is a necessary part of the grammar governing “to perceive” that “only I have access to my own perceptions”. This is a restriction on the proper use of “to perceive”. It would be inaccurate to say “My friend has my perception of green” or “I have my friend’s sense-impression of softness”, but these are trivial truths. More importantly, this grammatical remark *does not* allow us to conclude that the existence of the external world can be called into doubt.

The metaphysical solipsist is in bad shape as well. There are two initial arguments against her position, both relying on our previous conclusions. Firstly, the “I” in “I am in pain” does not refer in the typical sense of ‘reference’ and secondly “I am in pain” essentially means the same thing as groaning does and groaning does not pick out the Self the solipsist believes to exist.²⁰ If the solipsist insists that her position is coherent and

²⁰ This ‘essentially’ will be fleshed out later. The statement that “I am in pain” and groaning are exactly equivalent is perhaps too crude to be correct. Let us suspend judgment until later on.

that it is simply a deficiency in our language that prohibits her from expressing her point, we have to insist that the very expression of her point relies on an abuse of the meanings of the words she employs.

The first argument is best understood after making a distinction. The referential aspect of a statement like “I am in pain” differs from statements like, “I have ten fingers”. In “I am in pain” the statement is using the ‘subject-use’ of reference while in “I have ten fingers” “I” is functioning as an ‘object-use’ of reference. Wittgenstein endorses this distinction:

Now the idea that the real I lives in my body is connected with the peculiar grammar of the word “I”, and the misunderstandings this grammar is liable to give rise to. There are two different cases in the use of the word “I” (or “my”) which I might call “the use as object” and “the use as subject”. [...] One can point to the difference between these two categories by saying: The cases of the first category involve the recognition of a particular person, and there is in these cases the possibility of an error, [...] On the other hand, there is no question of recognizing a person when I say I have toothache. To ask “are you sure that it’s *you* who have pains?” would be nonsensical. (BB, p. 66-7)

In these cases of object-use reference, there is a possibility of misidentifying the true referent. These possibilities are not present in the subject-use sense of reference. We cannot mistakenly refer to ourselves when we say “I am in pain”. It makes no sense to say “I do not know whether or not I am in pain” (see PI 408). These considerations illustrate the difference between the referential aspect of the object-use of “I” and the meaning of subject-uses of “I”.

The solipsist might agree to this distinction and continue to insist that they are simply advocating a truth hitherto unacknowledged about the metaphysical structure of the world. They are trying to use our language to stress a new point, trying to overcome the limits of the ordinary uses of our words. Edward Minar acknowledges this point as well in his article “Wittgenstein on the Metaphysics of the Self: The Dialectic of Solipsism in the Philosophical Investigations”:

This Wittgensteinian line of thought, I believe, constitutes a considerable criticism. From the philosopher's perspective, however, it continues to ring hollow. For him, it is no surprise that the function of 'I' is different from that of paradigmatic referring expressions; why else does he start searching for the peculiar entity to which 'I' must refer? The solipsist agrees, indeed insists, that subject-uses of 'I' do not meet the conditions that would yield a rich notion of reference; they lead to no worldly entity, however special. Although 'I' does not, on the solipsist's view, pick out an individual of an established kind, it must direct attention to a special subject *position*. Presented with conditions on ordinary reference that his search for a self must apparently meet, and with reasons (in the world, as it were) that he cannot meet them, the solipsist will question whether he needs to be responsible to these requirements. (Minar 1998, p. 339-40)

This is a legitimate response of the solipsist. Perhaps she means something different than what she seems to mean when she says that "I" refers to a transcendental Self. However, we have built up a fair amount of machinery to deal more thoroughly with the solipsist.

Using our earlier results, let us analyze the meaning of the phrase "I am in pain". The meanings of this and other similar first person psychological avowals comprise the evidence for the metaphysical solipsist's position. If we can show that the meanings of these phrases do not provide any support for the solipsist, then their position will be stripped of its foundation and collapse. The first thing to note is that "I am in pain" is used in remarkably similar situations that groaning is, often occurring simultaneously. There is a sense in which groaning 'refers'. It refers in so far as it *draws attention* to a particular person. Minar discusses this difference as well:

How has the conclusion that the 'I' in 'I am in pain' does not point to a person been reached? It can hardly be denied that the use of the sentence calls attention to the person who utters the words and in this limited sense distinguishes him from others (see S 406). Further, the uttering of the words licenses the conclusion that *someone* is in pain (see S 407). These facts alone do not, however, evidence a self to which 'I' serves to refer. They do not, in particular differentiate saying 'I am in pain' from groaning. A groan will manage to call attention to someone, but clearly without naming or referring to that person. (Minar 1998, p. 337)

It is clear that the meaning of groaning does not refer to any kind of subject. Raising our hand, screaming, or doing anything to call attention to ourselves is not paradigmatic of singular referring terms. Recall that the meaningfulness of any expression is completely bound up in the circumstances and actions particular to its occurrence. Now the essential

line of argument is fairly straightforward: the meanings of both of these expressions are more or less equivalent, and yet we *know* that groaning has no trace of referring to a transcendental Self.²¹ Therefore, “I am in pain” does not refer to a Self either, thereby disarming metaphysical solipsism.

This conclusion faces some initial objections and so needs elaboration. The solipsist might want to insist that even though she cannot express her point in her behavior or language that she knows what she ‘means’ anyway. After all, surely only I have THIS pain! This kind of statement rings of the kind of privacy we have been fighting against. Private acts of mental ostensive definition do not establish the meanings of our words and actions. So if the solipsist insists that her phrase “Only my experiences are real” means something, then its meaningfulness cannot extend beyond the action coordinate with it. This means that it will have to, in principle, be able to be communicated to another person. The meaningfulness of any expression lies in its use, in its actual instantiation, and so in principle should be observable by anyone else. It is true that we allowed for a *kind* of private language in chapter three, but this language is not in principle shut off from the possibility of anyone else’s learning it. The essential part of the positive private language argument (PLA) was that someone could develop a language while isolated from a community of people, not that someone could have a language that no one could ever understand. The position criticized by the negative PLA

²¹ I say ‘more or less’ equivalent to acknowledge a possible source of slight divergence in meaning between the two expressions. “I am in pain” is a particular occurrence of the general form “I am X”, where “X” can be any psychological state. Groaning is not the instantiation of any general form, and so we might be inclined to believe that there is a difference in meaning between the two expressions. However, I argue that the mere grammatical fact of our language that allows us to talk about general forms of propositions does not give us genuine insight into the meanings of particular utterances of phrases like “I am in pain” in their actual uses. “I am in pain” and “I am happy” might be grammatically very similar in ways that groaning and cries of joy are not, but the actual uses of groaning and “I am in pain” *are* extremely similar, as are the uses of “I am happy” and cries of joy. *These* similarities are what is important for meaning.

allows for the possibility of such a language, but it is untenable due to the incoherencies that follow from its traditional assumption about meaning (TA). This brings our discussion full circle: the thesis the solipsist asserts lacks genuine meaning because they cannot 'mean' anything privately.

As in the case of the epistemological solipsist, the metaphysical solipsist has been tricked by the difference in grammar between legitimate cases of reference and the grammar governing first person avowals of psychological states. We need to reject the 'model of object and designation' when we consider the meaning of the word "I", for the self is *not* an object. ("I" is not the name of a person, [...] (PI 410)) Just as numbers and mental states are not objects, the self eludes ostensive definition. As in the case of mental states, we can meaningfully talk about the self despite the fact that it is not an 'object'.

This concludes the link between Wittgenstein's discussion of rules, meaning, private language, mental states, and the self, although it leaves us with decidedly negative conclusions. Wittgenstein wants to retain the meaningfulness of words "myself", "I", "self-knowledge" and the like, but he does not give very helpful ideas as to how this should be done. We are supposed to simply look at their employment and see the meaning there. The problem with this response is that we have been so far removed from the *real* meanings of phrases that use these words that assumptions like TA, CA, and SA have deeply permeated our thought. Dismissing these assumptions moves us in the right direction but leaves us short of genuine insight into the real meanings of words used to discuss 'the self'. While the later Wittgenstein advanced a kind of holism that we have been working to understand, this holism does not give us a positive account of the self. Rather, it gives us powerful reasons for dismissing assumptions thoroughly ingrained in

our conception of these problems, which is no small accomplishment. The epilogue that follows serves as my attempt to end on a positive note. There does seem to be genuine work that still needs to be done, even if we accept Wittgenstein's conclusions, and I take this opportunity to show where I think we should direct our efforts.

Epilogue:

Final Thoughts and Suggestions for Future Research

This essay has been concerned with following a line of thought through Wittgenstein's views on the meaningfulness of language to his discussion of solipsism and the self, but in doing so it also defends Wittgenstein's position as a tenable alternative to theories endorsing any combination of TA, CA, or SA. Wittgenstein's position gives us powerful tools against some traditionally overwhelming problems, namely the mind/body problem and the threat of solipsism. The best immediate upshot of the essay is that we do not have to spend our future efforts looking for 'a self' that is an object, much less how one privately has sensations in their minds. Neither sensations nor the selves are objects, and so it does not make sense to ask about how the self 'has' sensations. However, this and our other conclusions have been primarily negative. This epilogue addresses two questions: where should our future efforts be directed and is there any positive conclusions we might draw about ourselves from these meditations.

There are several areas of this essay that could be spring boards for further research. While tangential to the core of the arguments in this essay, the epistemological questions raised in the third chapter demand further exploration. The role of custom and the ensuing relativism between forms of life raises important questions about correcting others' uses of words. Perhaps the training each of us received differs to such a slight degree that we all slightly differ in our uses of words. We need to see how this might affect meaningful communication between members of the same language community. There also seems to be an important relationship between the notion of our highly complicated extended language games and art. This essay portrays the construction of

meaningful language as an exploration of a limitless horizon, where the value in our quest is up to us. Art seems to be one of the most frequented harbors from which we embark on these adventures, but the extent to which this metaphor is accurate will depend on greater thought and discussion. On a similar note, the relationship between ethics and our conclusions demands further attention. Our status as moral *agents* is most likely affected in some way by the denial that our selves are ‘objects’. Finally, the pursuit of a direct solution to Wittgenstein’s rule-following considerations should not be abandoned entirely. Frege and Russell’s responses, as portrayed here, do not appear adequate, and chapter two merely serves as a possible account of meaning that rejects the traditional assumptions about meaning that lead us into these problems. However, I do not claim that any direct solution is impossible *in principle*. I have merely argued that highly influential accounts fail and that we can hold on to a coherent theory of meaning while rejecting traditional assumptions.

When it comes to a more positive account of the self, Wittgenstein does not explicitly tell us what the meanings of words like “myself” and “I” are, much less words like “self-actualization” or “self-knowledge”. Yet these and related words are not just elements of a philosopher’s musings: they are bound up in *real* questions that pose serious problems for many people. We might admit that groaning and “I am in pain” have equivalent meanings, but what about questions like “What do *I* truly value?” or “Do *I* love this person?” The meaningfulness of these questions and their answers cannot be reduced to some series of primitive gestures, nor would Wittgenstein attempt such a reduction. A robust account of the self should not leave us in the dark about these sorts of questions. The focus of our philosophical inquiry should be directed at what self there is

that can be known, which in turn would show us how we can ‘know ourselves’ and what it means to know other people. We are looking for a theory that gives some genuine insight into what it is like to be a human being, not one plagued by misleading assumptions or one that stops short of answering these questions.²²

Despite falling short of this grander goal here, Wittgenstein’s thoughts might serve as the foundation for a new conception of the conditions under which we acquire genuine self-knowledge. We have seen that the meanings of our words are connected with the actions and circumstances surrounding their use. This conclusion affirms some generally accepted but often unappreciated truisms about the relationship between yourself and your community. It is frequently taken for granted that your friends and family have genuine insights into aspects of your character of which you are otherwise unaware. There is a temptation to dismiss the importance of this fact on the basis that, at least in principle, someone with enough self-awareness would realize the true nature of their self independently of their friends’ input. This belief is bolstered by the apparent fact that we each have privileged access to some kind of substantive private knowledge about ourselves. Furthermore, some might take these views to claim infallible knowledge about their self. Each of us alone knows who we really are, and the conjectures of others can only have probabilistic accuracy about my true character.

We have denied that there is a private self serving as an object that we can observe through private communion. To flesh out the errors of this conception, let us

²² Hans Sluga is sensitive to this issue in his article “‘Whose house is that?’ Wittgenstein on the self”. He considers the sentence, “I am not a genius, I am only a talent”, pointing out that the meaning of “I” in this case can neither be reduced to the meaning of a groan nor establish a self that is the referent. He goes on to argue for a constructivist conception of the self. We have the idea of a self as an object that serves as a necessary illusion with which we make sense of ourselves as moral and historical agents (see Sluga, p. 346-8). It is not clear to me that we are forced into such an account given Wittgenstein’s considerations, but it is worth mentioning that others have tried to extend Wittgenstein’s reflections on the self into a more robust theory.

return to an example mentioned in chapter two: that of the man who sincerely said he loved his wife even though he neglects and beats her. Suppose this man says of himself, “I am a loving person”. What are we to do if we take this man at his word? As before, we would probably say that he is *not a loving person*, though he is inclined to think himself one. If he insists that he is, in fact, very loving, our next move would be to check if he means the same thing we do by “to love”. This procedure might take many forms, but presumably it would be analogous to our discussion of the improper use of the word “green” in chapter three (although the meaning of “to love” might prove more difficult to teach). There are three options once we reach bedrock. (1) He might acknowledge that “to love” really means what we mean by “hateful” or “malicious” for him, in which case we can all at least agree on this point. The other two options are more confusing: either (2) he is not able to learn the meaning of “to love” or (3) he learns or demonstrates that he already knows how to use “to love” correctly and yet insists that he is a loving person. Case (2) would be bizarre, and might only be explainable by an appeal to a different form of life. The third case is more important for our purposes because no one speaks past each other in this case. Now our response would be the same as our initial one. The point here is that we are *right* in observing that he is not a loving person and he is *wrong* to think otherwise, no matter what special private feelings he may have. No matter how tenaciously he holds this belief, he will not be a loving person.

This example demonstrates a real possible scenario where others’ observations can be genuinely accurate about another person’s character despite the individual’s contrary avowals. The greater question remains: was it in principle impossible for this man to have realized the falsity of his statement without the observations of others? The

answer seems to be no. It seems perfectly possible that this man wakes up one day to a terrible epiphany. But this possibility does not prove that we have a privileged private access to ourselves that is in principle inaccessible for anyone else. In fact, we see that other people can even have *real* knowledge about yourself in spite of your believing otherwise.

These considerations derive their strength from the elusive grammar of words like “I”, “myself”, “yourself”, and so on. The meanings of these words are perfectly familiar in so far as we use them uniformly. Things become more confusing when we ‘try to understand ourselves’. We have been under the illusion that “I” refers to a subject that we alone have access to when no such object exists. Of course, “I am a loving person” is still a meaningful expression, as is “I am in pain”, but the grammar of these phrases does not preclude others from having genuine insights into your character or mental states. We fall prey to the belief that we know exactly who we are because of the complimentary notions of privacy and the metaphysics of the self. But if we reject these ideas, as we have above, we see that you can learn a great deal about yourself through regular open dialogue between yourself and those who know you. The advice and insight of your friends and family is not mere conjecture. They frequently see things that you do not and listening to them will give you opportunities for self-knowledge and growth that would be more difficult to arrive at otherwise. “Myself”, then, looks to be something that is genuinely accessible to others. In fact, Wittgenstein’s thoughts have brought us to a more powerful understanding of an old truth: we are “social beings”, and by this we mean that our peers have direct access to our ‘selves’ in the social sphere of our lives. Moreover, we come to know ourselves in the midst of our fellow humans, without whom we would be apt to

forget and/or misconstrue our past actions into ways that distort our knowledge of ourselves.

These final meditations reveal that much work lies ahead of us. We have the beginnings of a more positive account of the self and spring boards for further research. We owe a debt to Wittgenstein's insights, but it seems that we have many questions left unanswered by his account. While Wittgenstein's thoughts may not have extended this far, we can use what vision he had to build something stronger than the house of cards he blew over.

Bibliography

Works by Wittgenstein:

Wittgenstein, Ludwig. *Blue and Brown Books*. New York: Basil Blackwell, 1958.

Wittgenstein, Ludwig. *Philosophical investigations*. Upper Saddle River: Prentice-Hall, Inc, 1958.

Wittgenstein, Ludwig. *Remarks on the Foundations of Mathematics*. Cambridge: Basil Blackwell, 1978.

Wittgenstein, Ludwig. *Tractatus Logico-Philosophicus*. New York: 2nd ed. Routledge, 2001.

Wittgenstein, Ludwig. *Zettel*. Berkeley: University of California Press, 1967.

Works by Other Authors:

Baker, G. P. and P. M. S. Hacker. *Scepticism, Rules and Language*. Oxford: Basil Blackwell Publisher Limited, 1984.

Boghossian, Paul A. "The Rule-Following Considerations." *Mind* 392 (Oct., 1989): 507 – 49.

Boghossian, Paul A. "Rules, Meaning, and Intention." *Philosophical Studies* 124 (2005): 185-97.

Buchler, Justus. *Philosophical Writings of Peirce*. New York: Dover Publications, Inc, 1995.

Hacker, Peter. *Insight and Illusion*. Oxford: Clarendon, 1972.

Kripke, Saul. *Wittgenstein On Rules And Private Language*. Cambridge: Harvard University Press, 1982.

McDowell, John and Philip Pettit. *Subject, Thought, and Context*. Oxford: Clarendon Press, 1986.

McDowell, John. "Criteria, Defeasibility, and Knowledge." *Proceedings of the British Academy* LXVIII (1982): 455 – 80.

Minar, Edward H. "Wittgenstein on the Metaphysics of the Self: The dialectic of Solipsism in the Philosophical Investigations." *Pacific Philosophical Quarterly* 79 (1998): 329-54.

Putnam, Hilary. *Pragmatism: An Open Question*. Malden: Blackwell Publishers Inc, 1992.

Russell, Bertrand. *The Philosophy of Logical Atomism*. Chicago: Open Court Publishing Company, 1985.

Sluga, Hans and David G. Stern. *The Cambridge Companion to Wittgenstein*. Cambridge: Cambridge University Press, 1996.