

*A Critique of Saul Kripke's*  
**Wittgenstein on Rules and Private Language**

by

Chrysoula Gitsoulis

A dissertation submitted to the Graduate Faculty in Philosophy in the partial fulfillment  
of the requirements for the degree of Doctor of Philosophy,  
The City University of New York.

2009

© 2009

Chrysoula Gitsoulis

All Rights Reserved

This manuscript has been read and accepted for the  
Graduate Faculty in Philosophy in satisfaction of the  
dissertation requirement for the degree of Doctor of Philosophy.

Professor Jonathan Adler

Chair of Examining Committee

Signature

Date

Professor Iakovos Vasiliou

Executive Officer

Signature

Date

Professor Paul Horwich

Professor Stephen Neale

Professor Michael Levin

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

**Abstract**

A Critique of Saul Kripke's

*Wittgenstein on Rules and Private Language*

by

Chrysoula Gitsoulis

Advisor: Professor Paul Horwich

In *Wittgenstein on Rules and Private Language*, Saul Kripke presents a controversial skeptical argument, which he attributes to Wittgenstein's interlocutor in the *Philosophical Investigations* [PI]. The argument purports to show that there are no facts that correspond to what we mean by our words. Kripke maintains, moreover, that the conclusion of Wittgenstein's so-called private language argument is a corollary of results Wittgenstein establishes in §§137-202 of PI concerning the topic of following-a-rule, and not the conclusion of an independently developed argument in §§243ff of PI, as most commentators take it to be. In this work, I assess Kripke's skeptical argument both in its own right, and as an interpretation of the rule-following sections of PI. In its own right, I try to show that it is critically flawed. However, as an interpretation of the rule-following sections of PI, I try to show that it is essentially correct. I do this by showing that Kripke's interpretation squares with and supports the metaphilosophical framework developed by Wittgenstein in §§107-136 of PI, which immediately precedes his remarks on following-a-rule.

## A Critique of Saul Kripke's Wittgenstein on Rules and Private Language

### Preface

Almost three decades after its publication, Saul Kripke's intriguing and controversial *Wittgenstein on Rules and Private language* (1980) (henceforth WRPL) remains one of the most important and widely read commentaries on §§137-242 of Wittgenstein's *Philosophical Investigations* (henceforth PI).

Many factors account for the book's enduring popularity. One is that it is written in simple, non-technical terms, which has made it readily accessible to a wide and diverse audience. Another is that it presents a captivating argument, which appears to flow seamlessly from apparently true premises to a startling conclusion. The crucial factor accounting for the book's popularity, however, is the important contribution it has made both to exegetical studies on Wittgenstein and to contemporary debates concerning the metaphysics and epistemology of meaning and understanding.

As a piece of historical scholarship, WRPL offers an interpretation of Wittgenstein that is far from ordinary. It represents Wittgenstein as defending a set of unified and highly provocative claims about the nature of meaning and understanding. Unlike many expositions predating WRPL, which take Wittgenstein's central topic of concern in his discussion of meaning and understanding in PI to be his famous 'private language argument', and locate it in §§243ff of PI, Kripke takes his central topic of concern to be that of 'following a rule' in §§137-202. Kripke's interpretation of these sections not only

helped put the topic of following a rule into the philosophical limelight, but also tied together, in a novel and striking way, what seemed to be a series of dissociated remarks on rule-following and the attendant private language argument.

On Kripke's reading, the conclusion of the private language argument is a 'corollary' of results that Wittgenstein already establishes in his discussion of following a rule. More specifically, Kripke takes the conclusion of the private language argument to be a corollary of what he calls a 'skeptical solution' to a 'skeptical problem' concerning the topic of following a rule. Indeed, according to Kripke, the private language argument already reaches its conclusion by §202 of PI.

The '**skeptical problem**', which Kripke takes to be the central problem occupying Wittgenstein's attention in §§137-202, is: *Is there any fact about me that constitutes what I mean by my words?* The main character of WRPL, an imaginary meaning skeptic, examines a range of facts that might play the requisite constitutive role – and so solve the skeptical problem – and shows that they cannot do the job required of them. He then considers the proposal that my meaning what I do by my words is a primitive fact, and cannot be constituted by more basic facts, but rules this out as well. These considerations form the content of what Kripke calls the '**skeptical argument**'. On the assumption that he has examined all possibilities, he draws the (apparently paradoxical) '**skeptical conclusion**' that *there are no facts about me that correspond to my meaning what I do by my words.*

Faced with this seemingly outrageous result (undercutting, as it apparently does, some of the most fundamental assumptions we make about the notions of meaning and understanding), we naturally incline toward the view that there must be something wrong

with the argument leading to it: that it relies on some assumption that we can reject, or that it makes some fallacious step. That is, we incline toward *rejecting* the skeptical conclusion, and maintaining that there must, after all, be what Kripke calls a '**straight solution**' to the skeptical problem: some species of fact that constitutes my meaning what I do by my words, and hence that explains my applying them in the way that I do. Kripke, however, reads Wittgenstein as advocating what he calls a '**skeptical solution**': a response to the skeptical argument which *accepts* the skeptical conclusion but seeks to explain how we can rehabilitate talk of meaning in its light.

The skeptical problem, argument, and solution, have received a great deal of attention, not merely as interpretations of Wittgenstein's discussion of meaning and understanding, but also in their own right. Many contemporary philosophers of mind and language are familiar with the skeptical problem, and have tried to offer a straight solution to it. Kripke's book makes an important contribution to efforts to 'solve' it, by ushering into contemporary debates the most important insights he takes Wittgenstein to offer on it. Ultimately, on Kripke's reading of Wittgenstein, the skeptical problem has no straight solution. This result proved to be extremely contentious for many contemporary philosophers of mind and language, most of whom maintain that there is a straight solution to the skeptical problem. Not surprisingly, WRPL gave rise to a wave of controversy, which continues to this day.

Efforts to solve the problems arising from these debates have not been entirely disconnected from the historical Wittgenstein's writings; indeed, WRPL has prompted many contemporary philosophers of mind and language to revisit Wittgenstein's work to ferret out whatever insights he might have had on them. This, in turn, has helped *bridge*

*the gap* between Wittgenstein's discussion of meaning and understanding and contemporary discussions of these topics. We can think of WRPL as a stepping stone between these domains. If Wittgenstein's discussion of meaning and understanding is taken more seriously in philosophy of mind and language today than in the past, the credit is due in large part to WRPL.

Both of these factors – WRPL's departure from standard expositions in its reading of Wittgenstein, and the radical theses WRPL establishes in its own right -- are responsible for the book's enduring popularity. And, of course, enhancing the book's appeal is the fact that Kripke defends many of his central contentions with force and an abundance of insights.

Though WRPL, like its predecessor *Naming and Necessity* (1972), has been widely read, its content has not been taken up as favorably; indeed, the general consensus has been that the arguments of WRPL are flawed both in their own right and as an interpretation of Wittgenstein. Colin McGinn, for example, in the Preface to his book, *Wittgenstein on Meaning*, writes: 'When I first read Saul Kripke's *Wittgenstein on Rules and Private language*, it seemed to me to make clear sense, at last, of Wittgenstein's text... In the course of rereading Wittgenstein with Kripke's interpretation in mind I came, to my surprise, to have considerable doubts about the correctness of that interpretation ... At the same time, I began to think that the arguments Kripke develops were less impregnable than had first appeared to me.' Many commentators on WRPL have voiced similar remarks/reactions.

It seems to me, however, that the negative reception of WRPL rests on a number of critical textual misinterpretations. When I first read WRPL, I too thought the arguments

it presented were flawed both in themselves and as a reading of Wittgenstein. After a more careful reading of WRPL, however, and of the associated passages in PI, many of the inconsistencies that I thought were present in WRPL began to dissipate, and my appreciation of it as a reading of Wittgenstein intensified. I came to understand why the skeptical argument took the rather unusual form it did, and that together with the skeptical solution, it forms a very good interpretation of the associated passages in PI. I also came to realize that most commentators failed to see this because, in some fundamental way, they misrepresented the content of WRPL: quoting passages out of context, omitting or distorting important parts of arguments, interpreting key expressions in ways not intended by Kripke, creating dialogue that is not present in the text, frequently affirming their opinion by corroborating the interpretive errors of others, and, most importantly, failing to distinguish two characters in WRPL -- Kripke's skeptic (who plays the role of the interlocutor in PI), and Kripke's Wittgenstein (who plays the role of Wittgenstein) -- thereby putting words into the mouth of the wrong character. Given that I thought the book had many valuable insights to offer that were being lost in the tangle of interpretive errors, I thought it well worth attempting to clear them up. That task is taken up in Chapter 1, where I present a detailed exposition of the skeptical problem, argument, and solution. In Chapter 2, I evaluate the skeptical argument in its own right. And finally, in Chapters 3 and 4, I examine how well the skeptical argument and solution fare as an interpretation of Wittgenstein.

I take these tasks to be important for several reasons. To begin with, regardless of whether or to what extent the skeptical argument has its source in PI, it demands careful evaluation in its own right, since it threatens to deplete our concepts of 'meaning' and

‘understanding’ of content, and thereby to undermine our everyday folk-psychological practices that depend on these notions; and it is not apparent that the skeptical solution can do the needed resuscitative work.

It is important to note that when I say the skeptical argument demands careful evaluation ‘in its own right’, I do not mean that I will not rely on PI in interpreting it. For I do not think the task of assessing the skeptical argument can be properly executed without appealing to the sections of PI which it is intended to illuminate. In fact, persistent failures to do this have, I think, been mainly responsible for the skewed readings of the argument.

It is not only the work of evaluating the skeptical argument in its own right that will precipitate us into PI, of course, but also, and necessarily, the task of assessing how well WRPL fares as an interpretation of PI. I take this task to also be important, for several reasons. For one thing, I have found that assessing WRPL as a reading of PI helps bring some of the most difficult passages of Wittgenstein’s discussion of rule-following into sharp focus. It also helps us develop a greater appreciation for Wittgenstein’s perspective on the skeptical problem, to the extent that we can locate it in PI, and other problems of a similar nature -- a perspective which, I believe, contains many important morals.<sup>1</sup>

In WRPL, as very often in Wittgenstein’s works, a problem is presented and, quite naturally, erroneously interpreted. In WRPL, it is the skeptic (who plays the role of Wittgenstein’s interlocutor) who is guilty of the *false interpretation*, and, on the basis of this error, he generates what Wittgenstein would call a ‘*pseudo-problem*’, since it is based on a false interpretation. Repeated attempts are made to solve it, to no avail. The

---

<sup>1</sup> Here, I agree with Crispin Wright [1989a p. 157] that ‘there are still rich seams of philosophy, which we have failed to mine only by digging insufficiently deep’.

skeptic then concludes that the problem has no solution; indeed, his false interpretation causes him to draw an even stronger conclusion that appears to be paradoxical. At this point Kripke's Wittgenstein enters the scene, and offers a way out of the seeming paradox: a 'skeptical solution'. This involves describing actual linguistic practice, to help us see why no direct solution to the pseudo-problem is necessary or possible. The moral to be drawn is that the pseudo-problem is not in need of *solution*, but of *dissolution*.

I take this to be the central interpretive thesis of WRPL, a thesis that has been lost sight of in the literature, and I think it is *essentially* correct as an interpretation of Wittgenstein. In chapters 3 and 4, I offer my defense of this claim. Wittgenstein once wrote: 'A human being is *imprisoned* in a room, if the door is unlocked but opens inward; he, however, never gets the idea of *pulling* instead of pushing against it.' [RFM III, 37] In similar vein, I try to show that the reason the skeptical problem, as interpreted by the skeptic, baffles us and frustrates attempts to solve it, is that we have been pushing against a door that only opens by pulling it. Pushing against the door is trying to give it a *straight solution*. Pulling it is offering a *skeptical solution*. I try to bring this out in Chapter 4 by situating the skeptical problem, as interpreted by the skeptic, and the skeptical solution, into a larger framework concerning Wittgenstein's so-called 'metaphilosophy' (remarks concerning the practice of philosophy), that deals with his identification and treatment of pseudo-philosophical problems. Kripke maintains that the conclusion of Wittgenstein's private language argument is a corollary of results Wittgenstein establishes in §§137-202 of PI concerning the topic of following-a-rule, and not the conclusion of an independently developed argument in §§243ff of PI, as most commentators take it to be. I go a step further in Chapter 4 and try to show that

Wittgenstein's remarks on following-a-rule in §§137-202 of PI are a corollary of his 'metaphilosophical' remarks in §§107-136 of PI, which immediately precede his discussion of following-a-rule; they are, in other words, but one application/illustration of his philosophical method. I also defend the correctness of Kripke's interpretation of the rule-following remarks, which see Wittgenstein as developing a skeptical argument and solution, by showing that this interpretation squares with and supports the metaphilosophical framework to be presented. A crucial goal of this project is thus not only to clarify the relation between Wittgenstein's remarks on following-a-rule and the private language argument, but the *more fundamental* relation between Wittgenstein's philosophical method and his remarks on following-a-rule. If Kripke's discussion helped put the topic of following-a-rule into sharp focus, by linking it to the private language argument, then my hope is to sharpen that focus a bit further, so as to illuminate the fundamental role that Wittgenstein's metaphilosophy plays in his discussion of following-a-rule.

That, then, is a broad overview of the plan and motivation for this project. I turn now to more specific details of content.

### *Chapter 1*

In §§137-242 of PI, Kripke interprets Wittgenstein as wrestling with the following problem about meaning: *Is there any fact about me that constitutes what I mean by a word?* This problem is developed by Kripke in WRPL in terms of an example from arithmetic. Suppose, the example runs, that I am given a computation that I have never performed before – say, '68+57=?'. I then perform the computation, obtaining '125' as

my answer, and after checking my work, I am confident that ‘125’ is the correct answer. Now suppose, Kripke continues, a ‘bizarre skeptic’ comes along who questions my confidence. ‘What fact about you makes it the case that, as you used “+” in the past, the answer you meant or intended for “68+57” should have been 125 rather than, say, 5?’ he asks. [p. 8]<sup>2</sup> ‘Perhaps’, he continues, ‘as you used the term “+” in the past, you meant some nonstandard function’, which he calls ‘quus’, and defines as follows:

$$\begin{aligned} x \oplus y &= x+y \text{ if } x, y < 57 \\ &= 5 \text{ otherwise} \end{aligned}$$

If I meant *quus* by ‘+’ in the past, then the answer I intended for ‘68 + 57’ should have been ‘5’, so that, if I am to accord with my past intentions, the answer I should now give is ‘5’. What fact makes it the case (makes it true) that I meant *plus* and not *quus*? This is the skeptical problem, and in [section 1.1](#), I try to clarify in further detail the nature of this problem.

The skeptical argument developed by Kripke in Chapter 2 of WRPL proceeds by canvassing a range of candidate straight solutions to the skeptical problem. These include the following:

- [A] A ‘rule’ coming ‘before my mind’ that ‘tells me’ how to apply ‘+’ in all future cases. [p. 15-16, 22]
- [B] An introspectible qualitative state of consciousness, such as an image or sensation, present when performing additions. [p. 41-51]

---

<sup>2</sup> All bracketed page number references are to WRPL, unless otherwise indicated.

[C] A disposition to give the right answer, i.e., to say that had I been queried about any question of the form ‘ $x+y=?$ ’, I would have answered with the sum. [p. 22-40]

The body of the skeptical argument consists in the skeptic’s examination and rejection of these candidate straight solutions, along with the proposal that meaning so-and-so  $w$  is a ‘primitive’ fact.<sup>3</sup> [p. 41-42, 51-53] Assuming that these facts exhaust the possibilities, he concludes that there are *no* facts that correspond to my meaning *addition* rather than *quaddition* by ‘+’, or indeed anything by any word, since the conclusion generalizes to any word.

In section 1.2, I fill in the details of the skeptical argument, and the pivotal role that the so-called ‘justification constraint’ introduced by Kripke’s skeptic plays in his rejection of the facts considered. On the interpretation I offer, that constraint can be summed up as follows:

[J<sub>1</sub>] F justifies my applying word  $w$  thus and so, given what I mean by it, if and only if (1) I have *direct (non-inferential) access* to F, and (2) I can *deduce a priori* from F that I ought to apply  $w$  thus and so.

In Chapter 3 of WRPL, Kripke outlines what he calls a ‘**skeptical solution**’ to the skeptical problem. There are two components to the ‘skeptical solution’. On the one hand, unlike a ‘straight solution’, the skeptical solution is not intended to refute the

---

<sup>3</sup> Kripke defines this as a state ‘not to be assimilated with sensations or headaches or any “qualitative” state, nor to be assimilated to dispositions, but a state of a unique kind of its own’. [p. 51-53]

conclusion of the skeptical argument. What makes the position ‘skeptical’ is that it *accepts* the conclusion of the skeptical argument:

[SC] For any subject S, there is no fact about S that corresponds to his meaning *addition* by ‘+’ (or anything by any word).

What makes it a ‘solution’ is that it nevertheless defends our ordinary *practice* of ascribing meaning and understanding to each other.<sup>4</sup> According to the skeptical solution, our ordinary practice of ascribing meanings to each other is not undermined by [SC], for, contrary appearances notwithstanding, it is not jeopardized by the failure of any fact to justify our applications of words in the way ([J<sub>1</sub>]) demanded by the skeptic:

There are at least two delicate issues regarding the skeptical solution that stand in need of clarification. One issue concerns the very coherence of the skeptical conclusion [SC]. On the face of it, [SC] appears to be *self-undermining*: if there is no fact about S that his meaning *addition* by ‘+’ (or anything by any word) corresponds to, for any subject S, then how can anyone succeed in stating the skeptical conclusion without contradicting himself?

It all depends on how we interpret the skeptical conclusion. A number of interpretations have been offered in the literature; among them some of the following:

[a] (Radical skepticism) All words are meaningless.

[b] (Error theory) All meaning-ascriptions are false.

---

<sup>4</sup>I follow Scott Soames [1998 p. 313] in this interpretation of the terms ‘skeptical’ and ‘solution’.

- [c] (Expressivism) Truth/falsity and factuality/non-factuality is not properly predicable of meaning-ascriptions.
- [d] (Anti-individualism) Meaning-ascriptions cannot be analyzed in terms of facts about the subject *alone*.
- [e] (Anti-reductivism) Meaning-ascriptions cannot be analyzed in terms of any collection of natural, non-semantic facts.

In section [§1.3.1](#), I flesh out each of these interpretations. Among them, only [a] - [c] are self-defeating or incoherent -- we will look at some arguments that bring this out – and, though [d] and [e] are coherent, I show that, along with the other candidates, they lack crucial features that Kripke attributes to the skeptical conclusion. The target conclusion, I try to show, is [f].

- [f] There is no state of an agent, analyzable or non-analyzable/primitive, that satisfies [J<sub>1</sub>].

The second interpretive issue that stands in need of clarification concerns the relation between the skeptical solution and Kripke’s interpretation of Wittgenstein’s ‘private-language argument’. Kripke sees the conclusion of the private-language argument as a ‘corollary’ [p. 68] to the skeptical solution. But in what sense is it a ‘corollary’? In [§1.3.2](#), I flesh out the connection.

## Chapter 2

Kripke raises four objections against the dispositional theory [candidate [C] above], the most promising straight solution, and indeed the most ardently defended in the current literature. They may be summed up as follows:

- (1) **Finitude objection:** Since my actual dispositions are *finite*, it's not true that if queried about any two numbers I will come up with their sum, as the dispositional theory maintains, for some numbers are too large for my mind to grasp. But surely I can mean addition by '+' despite the fact that I don't have dispositions to handle huge numbers.
- (2) **Error objection:** Any meaning-constituting fact must be able to account for the fact that some people are disposed to make *mistakes*. But the dispositional theory cannot account for this fact: according to it, how one *should* apply '+' is how one *would* apply '+'.
- (3) **Circularity objection:** The dispositional theory cannot succeed in giving a non-circular description of the facts that constitute my meaning what I do by my words.
- (4) **Justification objection:** Any meaning-constituting fact must 'show how I am justified' [p. 11] in giving one answer rather than another to any particular addition query. I.e., any meaning-constituting fact must be such that, however in fact I am disposed to apply '+', it should 'tell me' [p. 24] how I *should* apply it. But dispositions do not have this property.

These objections have a common form (as will become apparent when we look at them in detail in chapter 2): they impose a specific constraint on the dispositional theory, and then reject it as a straight solution to the skeptical problem for failing to meet that constraint.

In chapter 2, I examine and critique each objection. The chapter is divided into four sections, corresponding to the four objections. The plan for each section is as follows.

- i. I begin by laying out the argument corresponding to the objection in question, so as to identify the constraint that the objection imposes on the dispositional theory.
- ii. I then argue that the argument's success depends on an *unwarranted assumption* Kripke makes about how the constraint is to be interpreted, and that no meaning-constituting fact should be required to meet the constraint thus interpreted.
- iii. Finally, I offer a more reasonable interpretation of the constraint in question, and try to show that a more refined version of the dispositional theory than that entertained by Kripke is invulnerable to the objection when the constraint is thus interpreted.

I develop this more refined version of the dispositional theory – '**constructive dispositionalism**', as I call it -- in piecemeal fashion throughout chapter 2, as we deal with each objection in turn.

### *Chapter 3*

In order to properly explain §§137-242 of PI, on which Kripke focuses his attention, we must situate them within the methodological framework in §§107-136 that precedes them. The main purpose of chapter 3 is bring this methodological framework to light, and thus to set the stage for chapter 4, where I defend the claim that the skeptical problem, as the skeptic interprets it, *does* emerge in §§137-242 as just another pseudo-philosophical problem, not in need of solution, but of dissolution. In chapter 3, I try to clarify Wittgenstein's means of identifying pseudo-philosophical problems, and his means of treating them, by focusing on his so-called 'meta-philosophical' remarks in §§107-136 of PI. There is an important analogy Wittgenstein develops in those sections between philosophy and medicine, which sees the practice of philosophy as a form of therapy, a method of healing. Uncovering the reaches of this analogy in Wittgenstein's work is one of the most fruitful and illuminating ways of explaining the aim and spirit of his philosophical method, and in particular, his diagnosis and treatment of pseudo-philosophical problems. I try to do so in chapter 3 by addressing the following questions, and defending the stated answers: §3.1: Symptoms: What, for Wittgenstein, are the symptoms that are in need of philosophical therapy? (Proposed answer: Conceptual tensions.) §3.2: Diagnosis: How, for Wittgenstein, are the symptoms generated? (Proposed answer: From the urge to generalize across different 'regions of language'.) §3.3: Treatment: What is the nature of Wittgenstein's therapy? (Proposed answer: Describing actual and exploring possible uses of language.)

#### *Chapter 4*

In this chapter, I apply the results of chapter 3 to the problem that forms Wittgenstein's focus of attention in §§137-242 of PI: What does understanding or meaning something by an expression 'consist in'? By placing a false interpretation on this problem, I try to show that the interlocutor (who plays the role of the skeptic in WRPL) constructs a pseudo-problem. This pseudo-problem, I maintain, can be identified with the skeptic's interpretation of the skeptical problem in WRPL. Thus, I try to show that the skeptical problem does emerge in PI, but as just another pseudo-problem, not in need of *solution*, but of *dissolution*. Throughout §§137-153 of PI, repeated attempts are made to solve this pseudo-problem, to no avail. It is mainly in these sections, I try to show, that we can locate the 'skeptical argument' of WRPL. In §§154-155, 179-183 of PI, Wittgenstein urges us to examine the kinds of circumstances in which we use the expressions that were falsely interpreted by the interlocutor (and led to the construction of a pseudo-problem), and the role these expressions have in our lives. It is in these sections, I try to show, that Wittgenstein presents the framework for a 'skeptical solution' -- a means of *dissolving* the pseudo-problem -- whose details he proceeds to fill in in §§184ff. Thus, the overarching aim of this chapter will be to show that the key ingredients of Kripke's interpretation of Wittgenstein can be traced back to PI, and are adequately supported by it.

## Acknowledgments

The thoughts recorded in this dissertation are the product of many years of reflection on themes from what I take to be one of the most important works of 20<sup>th</sup> century philosophy: Ludwig Wittgenstein's *Philosophical Investigations*.

There were a number of people who helped me along the way on this philosophical journey. My greatest debt is to my dissertation supervisor: Paul Horwich. Luck should have it that he would arrive at CUNY Graduate Center at the same time as I decided to write my thesis on Wittgenstein, and was searching for an advisor. I had the great pleasure of attending a number of courses with him, which helped set the stage for this work, by drawing my attention to, and helping me articulate, some of the central unresolved problems of Wittgenstein's *Philosophical Investigations* and Saul Kripke's *Wittgenstein on Rules and Private Language* that would form my focus of study here. Every conversation I had with him on those problems helped me clarify the position that I would eventually articulate here. In his classes on Wittgenstein, Prof. Horwich developed a line of interpretation toward which I was fully sympathetic, and has served as one of the guiding lights of this dissertation: that Wittgenstein's discussion of meaning can be properly understood only through the lens of his meta-philosophical framework. That framework was developed by Prof. Horwich in a clear and systematic way in class, and I would like to thank him for deepening my appreciation of its merit and importance in Wittgenstein's work. I am also extremely grateful to him for his time and patience in all my conversations with him, for forcing me to think harder and more clearly about the issues, for helping me weed out inconsistencies, misleading expressions, and other

careless errors in draft after draft of this work, and for his encouraging remarks on a variety of occasions. All of this was enormously beneficial.

CUNY Graduate Center has been a wonderful place to study philosophy, and I consider myself very fortunate to have been given the opportunity to pursue my doctoral studies here. I would like to thank all members of the department for making my years here fun and memorable ones. I am especially grateful to Michael Levin and Galen Strawson for reading, and giving me very helpful feedback on, the entire thesis. Michael Levin never ceased to amaze me by delivering superb comments at lightning speed! I am also grateful to both for their very encouraging remarks along the way. Thanks as well to Iakovos Vasiliou, Claudine Verheggen, and Nick Pappas for commenting on specific chapters, and to Stephen Neale, Jonathan Adler, and Stephen Grover for reading the entire thesis and raising excellent questions at the defense. Scott Soames, when he was a visitor to the Graduate Center, gave me very helpful and detailed feedback on a paper I wrote for his class that was turned, over time, into a chapter of this work, and Simon Blackburn took the time to discuss ideas I was developing in that paper with me when he too was a visitor to the Graduate Center. I would like to thank them both as well. And, of course, I would like to thank Saul Kripke, for offering an interpretation of the *Philosophical Investigations* that proved to be so intriguing to me, and rich in substance, that I made it the focus of this dissertation.

My interest in Wittgenstein's philosophical writings was sparked as an undergraduate in a class taught by Colin McGinn at Rutgers. I would like to thank him for his patient and good-natured way of dealing with my often confused questions. Had I not been brought to recognize the importance of Wittgenstein's *Philosophical Investigations*

through McGinn's stimulating discussions of it, I might never have ended up writing a dissertation on it; and surely the final product would be in much worse shape had I not, as a result of having been exposed as an undergraduate to it, had the benefit (and joy I should add) of years of meditation on it.

Teaching undergraduates at the *City University of New York*, throughout my years as a graduate student, has been a highly rewarding academic experience, and I benefited in immeasurable ways from it. Over the years, my students forced me to practice philosophy in a way that is simple, honest, clear, relevant, and engaging, and I thank them for that, if nevertheless I did not always succeed in meeting these standards. I have tried to here. I thank them as well for all the lively discussions we had, and for making my work more meaningful.

When someone or something proves to be a source of inspiration for me, I find that it always has a positive effect on my philosophical work, since for me, there is *no* divide between my personal and philosophical existence – they blend seamlessly into each other. I have been greatly inspired by my friend, Pavlos Paravantsos, so surely, his presence in my life has had a positive effect on this work. I would like to thank him for his love and affection, his spiritual guidance, and for all the great discussions we have had about life. They are some of my greatest treasures.

Above all, I would like to thank my Mother, Panagiota Valasoglou Gitsoulis, for her unfailing wisdom, love, and support. This dissertation is dedicated her, *with boundless love and admiration*, and to the memory of my Father, Nikolaos Gitsoulis.

Chrysoula Gitsoulis

Graduate Center, City University of New York

October, 2008

# CONTENT

## I. Evaluating *WRPL* in its Own Right

<b>1. Exposition of <i>WRPL</i></b> .....	<b>1</b>
1.1 The Skeptical Problem.....	1
1.2 The Skeptical Argument.....	10
1.3 The Skeptical Solution.....	28
1.3.1 Interpreting the Skeptical Conclusion.....	32
1.3.2 Kripke's Interpretation of the Private Language Argument.....	61
<b>2. Evaluating the Skeptic's Arguments against the Dispositional Theory of Meaning</b> .....	<b>74</b>
2.0 Introduction.....	74
2.1 Overcoming the Finitude Objection.....	76
2.1.1 The Finitude Argument.....	77
2.1.2 Critique of the Finitude Argument.....	82
2.2 Overcoming the Error Objection.....	96
2.2.1 The Error Argument.....	96
2.2.2 Critique of the Error Argument.....	98
2.3 Overcoming the Circularity Objection.....	109
2.3.1 The Circularity Argument.....	109
2.3.2 Critique of the Circularity Argument.....	112
2.4 Overcoming the Justification Objection.....	120
2.4.1 The Justification Argument.....	120
2.4.2 Critique of the Justification Argument.....	127
2.5 Conclusion.....	153

## II. Evaluating *WRPL* as a Reading of Wittgenstein

<b>3. Wittgenstein's Philosophical Method</b> .....	<b>156</b>
3.0 Introduction.....	156
3.1 Symptoms.....	157
3.1.1 Some Remarks on Style.....	165
3.2 Diagnosis.....	171
3.2.1 How Pseudo-Problems are Generated.....	171
3.2.2 The Myth of Platonism.....	178
3.2.3 The Moral Dimension of Wittgenstein's Philosophical Method.....	184
3.3 Treatment.....	190
3.3.1 Describing <i>Actual</i> Uses of Language.....	191

3.3.2	Exploring <i>Imaginary</i> Uses of Language.....	197
<b>4.</b>	<b>An Assessment of Kripke's Interpretation of Wittgenstein.....</b>	<b>206</b>
4.1	The Skeptical Problem of PI.....	206
4.2	The Skeptical Argument of PI.....	209
4.3	How False Interpretations Generate a Pseudo-Problem.....	215
4.4	The Skeptical Solution of PI: Dissolving the Pseudo-Problem..._	218
4.4.1	Dissolving the 'Inner State' Model of Meaning.....	218
4.4.2	Dissolving the Metaphysical Constraint in [J].....	223
4.4.3	Dissolving the Epistemological Constraint in [J].....	228
4.5	Conclusion.....	235
	<i>Bibliography</i> .....	237
	<i>Biographical Statement</i> .....	251

“Well, I assume that he acts as I have described.  
Explanations come to an end somewhere.”

Ludwig Wittgenstein  
*Philosophical Investigations*

# 1

## Exposition of *WRPL*

### CONTENT

---

1.1	The Skeptical Problem
1.2	The Skeptical Argument
1.3	The Skeptical Solution
1.3.1	Interpreting the Skeptical Conclusion
1.3.2	Kripke's Interpretation of the Private Language Argument

---

### 1.1 THE SKEPTICAL PROBLEM

In §§137-242 of the *Philosophical Investigations* [PI], Kripke interprets Wittgenstein as wrestling with the following problem about meaning: *Is there any fact about me that constitutes what I mean by a word?* This problem is developed by Kripke in *Wittgenstein on Rules and Private-language* [WRPL] in terms of an example from arithmetic. Suppose, the example runs, that I am given a computation that I have never performed before – say, ‘ $68+57=?$ ’. I then perform the computation, obtaining ‘125’ as my answer, and after checking my work, I am confident that ‘125’ is the correct answer. Now suppose, Kripke continues, a ‘bizarre skeptic’ comes along who questions my confidence. ‘What fact about you makes it the case that, as you used “+” in the past, the answer you

meant or intended for “68+57” should have been 125 rather than, say, 5?’ he asks. [p. 8]<sup>5</sup> ‘Perhaps’, he continues, ‘as you used the term “+” in the past, you meant some nonstandard function’, which he calls ‘quus’, and defines as follows:

$$\begin{aligned} x \oplus y &= x+y \text{ if } x, y < 57 \\ &= 5 \text{ otherwise} \end{aligned}$$

If I meant *quus* by ‘+’ in the past, then the answer I intended for ‘68 + 57’ should have been ‘5’, so that, if I am to accord with my past intentions, the answer I should now give is ‘5’. What fact makes it the case (makes it true) that I meant *plus* and not *quus*?<sup>6</sup>

Kripke’s skeptic proceeds by examining a range of facts that may be taken to solve this problem, and finds them inadequate. He then considers the proposal that meaning so-and-so by *w* is a primitive fact, and cannot be constituted by more basic facts, and finds this inadequate as well. Thus he draws the seemingly paradoxical conclusion that there is *no fact* about me that corresponds to my meaning *addition* by ‘+’, or indeed

---

<sup>5</sup> Kripke’s imaginary skeptic is merely a rhetorical device. The real problem is that of pointing out a fact that determines what I mean by a given word. In raising this problem, the skeptic is not raising a doubt about whether 125 is the correct answer, in the **arithmetical** sense; he grants that 125 is indeed, as a matter of arithmetical fact, the sum of 68 and 57. He is questioning what Kripke calls the ‘**metalinguistic**’ sense of correctness: What fact about me makes it the case that I used ‘+’ to denote the addition function? [p. 8] These two senses of correctness can come apart: if, e.g., I used the ‘+’ sign to stand for the subtraction function, 125 would still be the sum of 68 and 57, but the correct answer to the question ‘68+57 = ?’, in my idiolect, would now be ‘11’.

<sup>6</sup> Ridiculous and farfetched though it may seem, the skeptic’s hypothesis that I meant *quus* in the past, Kripke maintains, is **not logically impossible**:

To see this, assume the common sense hypothesis that by ‘+’ I *did* mean addition. Then it would be *possible*, though surprising, that under the influence of a momentary ‘high’, I should misinterpret all my past uses of the plus sign as symbolizing the quus function, and proceed, in conflict with my previous linguistic intentions, to compute 68 plus 57 as 5. ... The skeptic is proposing that I have made a mistake precisely of this kind, but with plus and quus reversed. [p. 9]

anything by any word. This conclusion appears to be paradoxical because, if there is no fact about any subject that corresponds to his meaning what he does by his words, then how can anyone succeed in stating the skeptical conclusion without contradicting himself?<sup>7</sup>

Kripke interprets §§137-242 of PI as developing a ‘skeptical paradox’ [p. 78] of this sort about meaning, and maintains that it requires a re-thinking of many otherwise plausible and tempting theories about meaning. [p. 31n] Wittgenstein, according to Kripke, tries to dissolve the paradox by offering what Kripke terms a ‘**skeptical solution**’ to the skeptical problem. [p. 79] According to this ‘solution’, Wittgenstein accepts the conclusion of the skeptical argument -- that there are no facts about me that correspond to my meaning what I do by my words -- but seeks to demonstrate that this result is not as radical or paradoxical as it might seem. Whether this interpretation of Wittgenstein is correct, or whether it misrepresents the character of Wittgenstein’s remarks in §§137-242, is an issue I will investigate in Chapter 4 of this project.

For future reference, it will be helpful to summarize the skeptical problem here:

[SKEP] Is there any fact about me that **constitutes**<sup>8</sup> my meaning *addition* by ‘addition’ (or, more generally, that constitutes my meaning *w* by ‘w’, for any expression *w* of our language)? [p. 11]<sup>9</sup>

---

<sup>7</sup> Whether the skeptical conclusion is indeed self-undermining in this sense is an issue I will examine in §1.3.1.

<sup>8</sup> Described as a ‘constitutive’ problem on e.g. pp. 11, 21, 39, 54, 69.

<sup>9</sup> Actually, the problem is presented, as I noted earlier, in the past tense: ‘Is there any fact about me that constitutes my *having meant* addition by “+”?’ But I will deal with it in its present tense form. One might wonder why Kripke presents the problem in the *past* tense. Kripke has the following to say: ‘We begin by speaking as if the notion that at present we mean a certain function by “plus” is unquestioned and

There are **two adequacy conditions** that the skeptic insists be satisfied by any ‘fact’ *F* that can be said to ‘constitute’ an agent’s meaning *addition* by ‘+’ (i.e., by any purported ‘straight solution’ to [SKEP]).

*1<sup>st</sup> Adequacy Condition on F*

*F* must be a non-relational ‘**state**’ of an agent that figures in an **analysis** (of any sort, reductive or non/reductive) of meaning addition by ‘+’. Talk of *F* being a ‘state’ suggests that *F* must be fully present/constructed/given; not potentially present/constructed/given. *F* can include not only physical states but also mental states or states of consciousness. Thus, Kripke says: ‘There are no limitations, in particular no behaviorist limitations, on the facts that may be cited to answer the skeptic. The evidence is not to be confined to what is available to an external observer, who can only observe my overt behavior but not my internal mental state.’ [p. 14]

*2<sup>nd</sup> Adequacy Condition on F*

*F* must ‘**contain**’ a **criterion of correctness**, a norm, by appeal to which a subject deduces that he ought to give a certain answer to a mathematical problem.

---

unquestionable. Only past usages are to be questioned, otherwise we will be unable to formulate our problem’. [pp. 13-14] There is a ‘deflated’ sense of ‘fact’ as it figures in meaning-ascriptions -- such as ‘It is a *fact* that Jones means *addition* by “+”’ -- in virtue of which Kripke’s remark is true, and a ‘robust’ sense in virtue of which it is false, as I will try to show in §1.3.1. Kripke’s remark, as I take it, is directed toward someone who seizes on the *deflated* sense, and builds it into his interpretation of the skeptical problem. Such a person might be baffled by the procedure, employed by the skeptic, of relying on the meanings of certain words in presenting the skeptical argument, while at the same time questioning what those words mean. The advantage of the formulation in terms of past usage is that it allows the skeptic to take the meanings of the words he uses in presenting the argument for granted, while the argument is being given and evaluated, thus removing this possible source of confusion (as noted by Scott Soames [1998 n.3, p. 342]).

[the skeptic] questions whether I have any **reason** to be so confident that now I should answer '125' rather than '5'. ... An answer to the skeptic must satisfy two conditions. First, it must give an account of what fact it is (about my mental state) that constitutes my meaning plus, not quus. But further, there is a condition that any putative candidate for such a fact must satisfy. It must, in some sense, **show how I am justified** in giving the answer '125' to '68 + 57'. The '**directions**' mentioned in the previous paragraph, that determine what I should do in each instance, **must somehow be 'contained' in any candidate for the fact as to what I meant**. Otherwise the skeptic has not been answered when he holds that my present response is arbitrary. [p. 11]

As a candidate for a 'fact' that determines what I mean, it fails to satisfy the basic condition on such a candidate, stressed above on pg. 11, that **it should tell me what I ought to do in each new instance**. [p. 24]

a fact as to what I mean now by '+' is supposed to **justify** my future actions, to make them inevitable if I wish to use words with the same meaning with which I used them before. This was our fundamental requirement on a fact as to what I meant. [p. 40]

This latter adequacy condition is composed of two sub-conditions. One sub-condition, which I will term the *metaphysical* constraint (concerning the *normativity* of meaning) on any purported meaning-constituting fact F, requires that F have a certain **nature**. The other sub-condition, which I will term the *epistemological* constraint on F, requires that F have a certain **role**.

**Metaphysical constraint** (concerning the *normativity* of meaning): F contains a norm, a standard of correctness, which determines in advance how w is to be applied in an open-ended range of situations.

**Epistemological constraint**: I have direct access to F (so that it can guide future applications of w).

We can summarize the metaphysical and epistemological constraints on any purported meaning-constituting fact F as follows:

[J<sub>I</sub>] F justifies my applying word w thus-and-so, given what I mean by it, if and only if (1) I have **direct (non-inferential) access** to F, and (2) I can **deduce a priori** from F that I ought to apply w thus-and-so.

This is what I take the justification requirement in the passage above to be. I will refer to [J<sub>I</sub>] as the skeptic's **justification constraint**. Thus we have:

[metaphysical constraint] + [epistemological constraint] = [justification constraint]

If F contains instructions on how to apply w in an open-ended range of situations (metaphysical constraint), and I have direct access to F (epistemological constraint), then I will be able to deduce a priori from F that I ought to apply w thus-and-so in any given situation, as [J<sub>I</sub>] requires.

[J<sub>I</sub>] is an **internalist** account of justification, which makes reference only to an individual. Condition (1) in [J<sub>I</sub>] (i.e., the epistemological constraint) insists that any purported meaning-constituting fact F be **readily available** to me, i.e., that I have **unmediated knowledge** of it, such as I do when I experience a pain, a tickle, a red after-image, or ringing in the ears. The subject has direct (non-inferential) access to such states because, 'since they are events in his consciousness, he is in the nature of the case

conscious of them'.<sup>10</sup> That this constraint be met by F is suggested in the quote from p. 24, where Kripke says that 'it should **tell me** what I ought to do in each new instance'. Further evidence for this condition can be found on p. 40:

The idea that we lack '**direct**' **access** to the facts whether we mean plus or quus is bizarre in any case. Do I not know directly, and with a fair degree of certainty, that I mean plus? Recall that a fact as to what I mean now is supposed to justify my future actions, to make them *inevitable* if I wish to use words with the same meaning with which I used them before. This was our **fundamental requirement** on a fact as to what I meant.

That the skeptic insists, in addition, that any purported meaning-constituting satisfy the metaphysical constraint is suggested in the quote from p. 11, where in talking about 'directions I gave myself', he says these directions 'must be **contained** in any candidate for the fact as to what I meant'. This requirement is also suggested in other places, where the skeptic insists that any putative meaning-constituting fact 'compel', 'establish' [p. 13, 21], 'show' [p. 41], 'entail' [p. 53], 'mandate' [p. 57], 'determine in advance' [p. 65], 'guarantee' [p. 95] that I mean *addition* by '+'.

The fact F that we are looking for must, on this interpretation of the justification constraint, be such that it provides the basis for an a priori deductive inference to claims about what I mean, and be transparent to me. Under these conditions, F will give me a **reason** for thinking '125' is the answer I ought to give, as the skeptic demands. [p. 11] I will then be able to cite F as a *justification* for the answer I give in order to show it isn't just a shot in the dark, 'a mere jack-in-the-box unjustified and arbitrary response'. [p. 23] This is why Kripke sometimes [see e.g. pp. 11-13] expresses the skeptical problem in

---

<sup>10</sup> As Crispin Wright observes [2001 p. 149].

*epistemological* terms: *How do I know* that I now mean by ‘+’ the same thing that I meant in the past by ‘+’? In answering this question I offer a *reason*: ‘I know *because ...*’

We need to distinguish this (strong) epistemological interpretation of the skeptical problem – [SKEP<sub>s</sub>] – which requires that any purported meaning-constituting fact F satisfy [J<sub>I</sub>], from a (weaker) non-epistemological interpretation -- [SKEP<sub>w</sub>] -- which does not require this. [SKEP<sub>w</sub>] may, e.g., require that F only satisfy the first adequacy condition – that, e.g., it serve as a *causal (reductive) explanation* of my meaning what I do by my words. A causal explanation is one that might not be immediately available to me; we use such explanations to explain, for example, why grass is green, or water is made of H<sub>2</sub>O. But the skeptic insists that any meaning-constituting fact be such that I have *immediate (non-inferential) knowledge* of it, and am able to *deduce* from it that I ought to apply w thus-and-so, given what I mean by w. So we must not confuse [SKEP<sub>s</sub>] with [SKEP<sub>w</sub>]. If there is a fact F that satisfies [SKEP<sub>w</sub>], and I have *immediate knowledge* of F, then I could appeal to my knowledge of F in attempt to offer a solution to [SKEP<sub>s</sub>]; however, any fact that solves [SKEP<sub>w</sub>] is not (*without further assumptions -- for example, assumptions about my knowledge acquiring powers*) *necessarily* one that will solve [SKEP<sub>s</sub>], for the constitutive fact might, e.g., be one that I am *not consciously aware of* – it might, for example, only be *tacitly* known by me. [SKEP<sub>w</sub>] concerns the *underlying nature* of meaning and understanding, and this underlying nature might not be *readily available* to me. And, even if it is readily available to me, I might not be able to *deduce* from it that I ought to apply w thus-and-so, given what I mean by w.

There is a *family* of expressions one could use in place of ‘meaning *addition* by “+”’ to formulate [SKEP]. For example:

understanding the term ‘+’

knowing the meaning of ‘+’

grasping the sense of ‘+’

mastering the concept expressed by ‘+’

following the ‘+’ rule<sup>11</sup>

So no straight solution to [SKEP] should be expressed in terms of members of this family (e.g. My meaning addition by ‘+’ consists in my *grasping the sense of ‘+’*) since one can simply reformulate the skeptical problem in terms of that member of the family (e.g. Is there any *fact* about me that my *grasping the sense of ‘+’* consists in?).

---

<sup>11</sup> Kripke often formulates the skeptical problem in terms of ‘following a rule’. It is customarily assumed that an individual’s words mean what they do in virtue of the individual’s following a rule for the use of those words: a rule that distinguishes correct from incorrect applications of the words.

We must beware of adding ‘intended *addition* by “+”’ to this list. As George Wilson points out in ‘On Skepticism about Rule-Following in Kripke’s Wittgenstein’ (forthcoming), from (1) ‘In using “+”, the speaker intended to be adding’, it doesn’t follow that (2) ‘The term “+” as the speaker used it meant *addition*’. For ‘it could be that the operation that the speaker learned to perform and did, in fact, perform was quaddition. That is, the speaker wrongly supposes that what he had learned to do is the same as the procedure or operation that others call “addition”’. The speaker could fail to do what he intended to be doing in connection with his use of “+”. If so, then “+”, as he was disposed to use it, actually meant “quaddition”, despite his own description of what he was doing as “addition”. ... What the skeptic questions is whether there could be a fact that constituted the speaker’s **succeeding** in what he then intended.’

As I like to formulate this point, given that [SKEP] is a problem about understanding, it should be distinguished from the more general problem:

[SKEP\*]    Is there any fact about me that constitutes what I mean by ‘+’ (or, more generally, by any word)?

If there is some fact about me that constitutes what I mean by ‘+’, then that same fact need not constitute my meaning addition by ‘+’, for I might mean quaddition by ‘+’. Thus, any solution to [SKEP\*] need not be a solution to [SKEP]; though of course we can’t solve [SKEP] unless we can solve [SKEP\*]. So sometimes in the ensuing discussion (next section) candidate solutions will be disqualified because of their failure to satisfy [SKEP\*]

## 1.2 THE SKEPTICAL ARGUMENT

The skeptical argument developed by Kripke in Chapter 2 of WRPL proceeds by exploring a range of candidate meaning-constituting facts, and then the possibility that meaning so-and-so by  $w$  is primitive, and cannot be constituted by any more basic facts. Taking it that the alternatives considered exhaust the possibilities, and fail to satisfy  $[J_1]$ , the skeptic concludes that there is no state that corresponds to meaning *addition* by '+', or indeed anything by any word.

The argument can be summarized as follows, the details of which I will return to shortly:

### *The skeptical argument*

i. If there is a state that corresponds to meaning *addition* by '+', then either it **can be analyzed** as:

[A] A 'rule' coming 'before my mind' that 'tells me' how to apply '+' in all future cases. [p. 15-16, 22]

[B] An introspectible qualitative state of consciousness, such as an image or sensation, present when performing additions. [p. 41-51]

[C] A disposition to give the right answer, i.e., to say that had I been queried about any question of the form ' $x+y=?$ ', I would have answered with the sum. [p. 22-40]

Or,

[D] Meaning addition by ‘+’ is a state that **cannot be analyzed**, i.e. it is a ‘primitive state’<sup>12</sup>. [p. 41-42, 51-53]

ii. States of type [A] – [D] do not satisfy [J<sub>I</sub>].

iii. Any state that corresponds to meaning *addition* by ‘+’ must satisfy [J<sub>I</sub>].  
(justification constraint)

iv. Therefore, states of type [A] – [D] do not correspond to meaning *addition* by ‘+’. [p. 13, 21]

v. Therefore (since we have exhausted all possible candidates), there is *no* state that corresponds to meaning *addition* by ‘+’ (and in particular, no meaning-constituting fact).

The use of the mathematical example, Kripke notes, is irrelevant to the presentation of the skeptical problem: the problem can be run on *all* uses of language.<sup>13</sup> Thus, the

---

<sup>12</sup> Kripke defines this as a state ‘not to be assimilated with sensations or headaches or any “qualitative” state, nor to be assimilated to dispositions, but a state of a unique kind of its own’. [p. 51-53]

<sup>13</sup> Kripke writes on page 19:

Of course these problems apply throughout language and are not confined to mathematical examples, though it is with mathematical examples that they can be most smoothly brought out. I think that I have learned the term ‘table’ in such a way that it will apply to indefinitely many future items. So I can apply the term to a new situation, say when I enter the Eiffel Tower for the first time and see a table at the base. Can I answer the skeptic who supposes that by ‘table’ in the past I meant *tabair*, where a ‘tabair’ is anything that is a table not found at the base of the Eiffel Tower, or a chair found there? Did I think explicitly of the Eiffel Tower when I first ‘grasped the concept of’ a table, gave myself directions for what I meant by ‘table’? And even if I did think of the Tower, cannot any directions I gave myself mentioning it be reinterpreted compatibly with the skeptic’s hypothesis?

skeptic concludes: ‘There can be no fact as to what I meant by “plus”, or any other word at any time’. [p. 21]

The reasons the skeptic offers for rejecting candidates [A] – [D] run as follows:

*A. A ‘rule’ coming before my mind*

The sought after fact cannot be a ‘rule’ coming ‘before my mind’ that ‘instructs’ or ‘tells’ me how to apply ‘+’. [p. 15-16, 22] Such a rule might take the form: Apply w to x iff x is so-and-so, where the so-and-so involves a **definition** of w in simpler terms. For example, the rule for addition might be given as follows:

- (1) Give answer z to  $x+y=?$  iff  $\langle x,y,z \rangle$  is in the extension of ‘+’, where the extension of ‘+’ is defined by the following (recursive) clauses:
- i.  $(\forall x)(x+0=x)$
  - ii.  $(\forall x)(\forall y)(x+s(y)=s(x+y))$ , where ‘s’ stands for the successor operation

A rule for addition might also take the form of a **recipe** or ‘**set of directions**’ [p. 15] (not necessarily involving a definition) for applying w. For example, a recipe for applying ‘+’ might be given as follows.

- (2) In figuring out the answer to the query ‘ $x+y=?$ ’, first count a heap of  $x$  marbles, then a heap of  $y$  marbles, and then count the union of the heaps thus formed. The result is the answer to ‘ $x+y=?$ ’

The problem that the skeptic points out for these proposals is that they only shift the skeptical problem to the *expressions that figure in the rule*. This invites further questions about what the meaning of the words occurring in the rule consist in.

For example, ‘count’ is one such term in the case of (2). The skeptic will ask: ‘What fact about you constitutes your meaning *count* by “count”? Maybe by “count” you mean “quount”, where to “quount” a heap is to count it in the ordinary sense, unless the heap is formed as the union of two heaps, one of which has 57 or more items, in which case you must give the answer “5”’. Since I applied the more basic concept (‘count’) in only finitely many past cases, it too can be interpreted in a non-standard quous-like way. [p. 16]

A similar problem emerges for (1). Here, ‘+’ has been defined in terms of ‘successor’, ‘universal quantification’, and ‘equality’. The problem is that these other functions, quantifiers, predicates, etc. have themselves been applied in only a finite number of instances. Thus, they too can be given non-standard interpretations, and so the skeptical problem surfaces for them all over again. E.g., ‘ $(\forall x)$ ’ might mean ‘for every  $x < h$  where  $h$  is some upper bound to the instances where universal quantification has hitherto been applied’, and similarly for successor and equality. [p. 16-17 n]

In sum, though I might have direct (non-inferential) access to any rule coming ‘before my mind’, the expressions that figure in that rule are open to further interpretation, so that

I cannot *deduce a priori* from it that I ought to apply *w* thus-and-so.<sup>14</sup> This shows that candidate [A] fails to satisfy condition (2) in [J<sub>1</sub>], and hence cannot serve as a straight solution.<sup>15</sup> From the standpoint of the skeptic, we are no better off than we were without the introduction of a rule. As Kripke says, ‘To say that there is a general rule in my mind that **tells me** how to add in the future is only to throw the problem back on to other rules that also seem to be given only in terms of finitely many cases.’ [p. 22]<sup>16</sup>

### ***B. Introspectible qualitative states of consciousness***

Perhaps it is not mental rules, but mental images and sensations that are the constituents of meaning and understanding, as candidate [B] suggests. For example, perhaps my meaning *cube* by ‘cube’ consists in bringing ‘before my mind’ a mental image of a

---

<sup>14</sup> I am not assuming that the ‘rule’ has its content built into it, so to speak, but has the status of string of symbols that stands in need of interpretation.

<sup>15</sup> Note that even if we removed the justification constraint, candidate [A] would still be vulnerable to the following **regress** objection:

[REG] How can invoking a rule R<sub>1</sub>— say, one that ‘comes before your mind’ — constitute your understanding a term? Given that R<sub>1</sub> will involve other terms, we can simply raise the skeptical problem again for those terms. If your understanding of one of those terms consists in a rule R<sub>2</sub> coming before your mind, then we can raise the skeptical problem once again for the terms involved in that rule, and so on, thereby generating a regress.

<sup>16</sup> Many commentators (e.g., Scott Soames [1997 p. 214] and Paul Boghossain [1989 p. 511]) have argued that this objection can be extended to mental content, i.e., to beliefs and other intentional attitudes. Since intentional attitudes have content, their content is as subject to the skeptical challenge as is the content of ‘+’. What, e.g., makes it true that an individual’s thought that ‘2+2=4’ is a thought about plus, and not a thought about quus? Whether the skeptical argument can be extended in this way is a controversial point, and not one that I have the space to examine here.

cube<sup>17</sup>, or in my having a special feeling whenever I use the word. Kripke considers this proposal on pp. 41-51.

Why not argue that ‘meaning addition by plus’ denotes **an irreducible experience, with its own special *quale*, known directly to each of us by introspection? (Headaches, tickles, nausea are examples of inner states with such qualia).** ... I referred to an *introspectible experience* because, since each of us knows immediately and with fair certainty that he means addition by ‘plus’, presumably the view in question assumes we know this in the same way we know that we have headaches -- by attending to the ‘qualitative’ character of our own experiences. Presumably the experience of *meaning addition* has its own irreducible quality, as does that of feeling a headache. The fact that I mean addition by ‘plus’ is to be identified with my possession of an experience of this quality. [p. 41]

Once again, the problem with this proposal is that it fails to satisfy clause (2) of [J]. The *quale* does not by itself determine the correct use of the associated word, because it is ‘semantically inert’<sup>18</sup>, i.e., stands in need of interpretation.

Suppose, for example, someone says that to mean *green* by ‘green’ is to bring up a color image of green whenever I use the word ‘green’. Cannot that image be interpreted in a grue-like way, so that it directs me to apply the word ‘green’ to *grue* objects? As Kripke says:

No internal impression, with a *quale*, could possibly **tell me** in itself how it is to be applied in future cases. Nor can any pile up of such impressions, thought of as rules for interpreting rules, do the job. The answer to the sceptic’s problem, “What tells me how I am to apply a given rule in a new case?”, must come from something outside any images or ‘qualitative’ mental states. [p. 43]

---

<sup>17</sup> This idea was made popular in the work of some of the British empiricists. For example, John Locke, in his *Essays Concerning Human Understanding*, associates the meaning one attaches to a word with a corresponding ‘mental idea’ or ‘mental image’, such as occurs in remembering, imagining, and dreaming. Kripke also associates it with Hume [See p. 63 n. 51].

<sup>18</sup> McGinn’s expression [1984 p. 7].

Not only do such contents fail to satisfy [J<sub>1</sub>], but, as Kripke further points out, they are *not necessary* for determining what I mean. [p. 42]<sup>19</sup> For, as a matter of empirical fact, mental entities do not always come ‘before our mind’ when we mean something by a word. Introspection does not bear out this assumption. [p. 42-43] Moreover, even when mental entities do come before our mind, it is not always the *same* mental entity. For example, when I use the word ‘red’, I might bring up the image of a red dress, a red apple, or a red flower, etc.<sup>20</sup> It is no more necessary for understanding an expression that I have a mental entity come before my mind – say, a mental picture of a cube whenever I hear or say the word ‘cube’ -- than it is that I have a concrete physical picture before me – a drawing of a cube.

We can actually strengthen this claim<sup>21</sup>: even in cases where there does seem to be an *empirical regularity* between a particular expression and a particular such item, we can still *conceive* of someone understanding an expression in the absence of that item.

So it looks like no kind of item coming ‘before the mind’ can be what meaning *addition* by ‘+’ consists in.

### ***C. The Dispositional Theory***

---

<sup>19</sup> This objection can be (and is on p. 23 of WRPL) applied to candidate [A] as well.

<sup>20</sup> Historically, Berkeley (*The Principles of Human Knowledge*) was one of the first figures to raise both of these objections in his critique of Locke’s theory of understanding.

<sup>21</sup> As Colin McGinn notes [1984 p. 6].

Perhaps we have failed thus far to produce the sought after meaning-constituting fact because we have been looking in the wrong domain for the constituents of meaning and understanding. Perhaps what constitutes a subject's meaning *addition* by '+' resides not in his mind, but in his overt physical behavior -- in the *applications* he makes of '+'.

It can't, however, be the *actual* applications a subject has made of '+' that constitutes what he means by it, for the skeptic's response to this proposal is all too obvious: alternative meanings are consistent with the actual applications that have so far been made of '+'. A subject's actual applications of '+' are *finite* in number, and it is consistent with his finite pool of answers to questions of the form ' $m+n = ?$ ' that he means something other than *addition* by '+'. Even assuming that a speaker has used '+' correctly in all his past uses, there is more than one meaning one can assign to '+' that is compatible with that usage: no finite selection of answers determines to within uniqueness what (if anything) the speaker means.

There is an obvious way one might attempt to surmount this difficulty. It is not *actual* behavior that constitutes what one means by a term, one may argue, but *un-actualized* behavior as well; that is, my meaning *addition* by '+' is determined by how I am *disposed* to apply '+'.

Kripke turns to this proposal on p. 22, where he considers the following **dispositional analysis** of meaning *addition* by '+':

[KD<sup>plus</sup>] A subject's meaning *addition* by '+' consists in the fact that for any p, q, he is disposed to answer r to the query ' $p+q=?$ ' iff r is the *sum* of p and q.

This proposal appears to be invulnerable to the objections that were raised against [A] and [B], since how a subject is disposed to apply a term is not a matter of some entity coming ‘before his mind’. In addition, it appears to surmount the underdetermination objection for actual behavior involving a term, since how a subject is disposed to apply a term involves more than his *actual* behavior involving that term: he might have been disposed to apply a term in a certain way even though he did not actually apply it in that way, if the circumstances appropriate for manifesting the disposition did not present themselves. In particular, it could be that he was all along disposed to answer ‘125’ to the question ‘ $57+68 = ?$ ’, but never actually did so, simply because no events occurred to trigger his additive disposition in that particular way.

Thus, this looks like a promising candidate indeed; one that is not easily demolished. Not surprisingly, Kripke devotes the bulk of Chapter 2 to that task. Whether he has succeeded is an issue I will take up in the next chapter.

Here, in brief, are the three objections that Kripke deploys against the dispositional theory.

*i. The finiteness objection* [pp.26-27]

As noted, the dispositional analysis attempts to surmount the underdetermination objection against actual behavior, but it doesn’t get very far, Kripke argues, for it is not true that if queried about the sum of *any* two numbers, no matter how large, I will answer with their sum: some pairs of numbers are simply too large for my mind to grasp. [p. 26] In this sense, my additive dispositions are *finite* – they can only extend so far. But surely

I can mean *addition* by ‘+’ despite the fact that I don’t have dispositions to handle huge sums.

This objection shows that the *right to left direction* of the analysis in  $[KD^{\text{plus}}]$  fails, i.e., that:

$\forall p, q, S$  is disposed to answer  $r$  to the query ‘ $p+q=?$ ’  $\leftrightarrow r$  is the *sum* of  $p$  and  $q$ .

It will not help the dispositional theorist to reformulate  $[KD^{\text{plus}}]$ , Kripke adds, by appealing to how I would respond under *ideal conditions* that overcome finiteness:

$[KD^{\text{plus}}]$  To mean *addition* by ‘+’ is to be disposed, *under ideal conditions*, when asked for any sum ‘ $x+y$ ’, to give the *sum* of  $x$  and  $y$  as answer.

because any characterization of the ideal conditions will ‘presuppose a prior notion of my having an intention to mean one function rather than another by “+”’. [p. 27] Suppose, for example, that the idealized conditions are described as ‘conditions where my brain is operated on in such a way that it would enable me to add, as opposed to quad, huge numbers’. It is evident that this idealization is circular, since it is characterized in a way that assumes I am disposed to add and not quad.

ii. *The error objection* [p. 29]

People can be disposed to make *mistakes*, so that the answer one *should* give, if it is to accord with the table of the function one actually means, can come apart from the answer

one is *disposed* to give. But, Kripke argues, a dispositionalist cannot say this: ‘according to him, the function one means is to be **read off** from his dispositions’. [p. 29] This, Kripke maintains, is because, according to the dispositional theorist, the value I mean ‘+’ to have for m and n is, *by definition*, the answer I would give if queried about ‘m+n’. [p. 25] The dispositionalist cannot *presuppose in advance* that a subject means plus and makes mistakes as opposed to meaning something non-standard that accords with the ‘mistakes’. For what he means is ‘read off’ the answers he is disposed to give. And so a dispositionalist must say that whatever answer a subject is disposed to give *is* right (since the function he means *is* the function that is ‘read off’ his dispositions). In order to avoid this consequence, Kripke maintains, the meaning-constituting fact must be distinct from the speaker’s dispositions. As Kripke puts it, a candidate for a fact must tell us what one *should* say, not what one *would* say. [p. 24] The dispositional view eclipses this fundamental **competence/performance distinction** by equating performance with correctness. [p. 24]

The error objection has the same basic form as the finitude objection, but, unlike the finitude objection, which shows that the right to left direction of the analysans in [KD<sup>plus</sup>] fails, the error objection purports to show that *both directions* of the analysans in [KD<sup>plus</sup>] fail, i.e. that:

S is disposed to give answer r to the query ‘p+q=?’ ↔ r is the *sum* of p and q

... ↔ ...

There are two ways a dispositionalist can respond to the error objection. He can either accept the claim that a subject who makes mistakes -- such as forgetting to carry -- does not mean *addition* by '+', but something non-standard. Or he can stick with common sense and say that a subject can mean *addition* by '+', even though his dispositions to calculate with '+' may not correspond perfectly to the addition function. In the latter case, according to Kripke, he would have to give up  $[KD^{\text{plus}}]$  as a way of formulating his theory.

Once again, Kripke maintains that it will not help the dispositionalist to reformulate  $[KD^{\text{plus}}]$  by introducing *idealized conditions* into  $[KD^{\text{plus}}]$ , as in:

$[KD_1^{\text{plus}}]$  To mean *addition* by '+' is to be disposed, *under ideal conditions*, when asked for any sum 'x+y', to give the *sum* of x and y as answer.

For he is then faced with the challenge (as in the finitude objection) of specifying the ideal conditions in a *non-circular* way, and, Kripke claims, he cannot succeed at this task.

*iii. The justification objection* [pp. 23-24]

The justification objection raises a more fundamental problem for  $[KD^{\text{plus}}]$  from the standpoint of the skeptic. The finitude and error objections show that my dispositions are not infallible, and do not cover the infinitely many cases of the addition function, so that satisfaction of the analysans in  $[KD^{\text{plus}}]$  is *not necessary* for meaning *addition* by '+'. The justification objection maintains that even if my dispositions *do* correspond perfectly

to the addition function (and hence the analysis is satisfied), that would not be enough to show that I mean *addition* by '+'. [p. 57]

So it does seem that a dispositional account misconceives the sceptic's problem – to find a past fact that *justifies* my present response. As a candidate for a 'fact' that determines what I mean, it fails to satisfy the basic condition on such a candidate, stressed above on p. 11, that it should *tell me* what I ought to do in each new instance. [p. 24]

Recall the second adequacy condition on any straight solution to [SKEP]: the skeptic insists that any candidate that constitutes my meaning one function, rather than another, by a given function sign, be such that, whatever answer I am in fact disposed to give, it should 'tell me' what answer I **ought** to give, and facts about how I am disposed to apply a term (even if in accord with its extension) do not, he insists, have this property. All that it seems can be recovered from the way in which I am disposed to use a term are claims about how I habitually do, will or would use it; not claims about how I *should* use it. So again, such facts fail to satisfy clause (2) of [J<sub>1</sub>].

There is another potential problem for the dispositional theory -- which infects physical theories of meaning and understanding more generally -- from the standpoint of the skeptic. The following passage suggests that Kripke identifies linguistic dispositions with physical states of the brain:

Am I supposed to justify my present belief that I meant addition, not quaddition, and hence should answer '125', in terms of a hypothesis about my past dispositions? (Do I record and investigate the past physiology of my brain?) Why am I so sure that one particular hypothesis of this kind is correct, when all my past thoughts could be construed so that I meant plus or so that I meant quus? [p. 23]

This identification brings out another problem:

**The idea that we lack ‘direct’ access to the facts whether we mean plus or quus is bizarre in any case. Do I not know directly, and with a fair degree of certainty, that I mean plus?** Recall that a fact as to what I mean now is supposed to justify my future actions, to make them *inevitable* if I wish to use words with the same meaning with which I used them before. This was our fundamental requirement on a fact as to what I meant. No ‘hypothetical’ state could satisfy such a requirement: If I can only form hypotheses as to whether I now mean plus or quus, if the truth of the matter is buried deep in my unconscious and can only be posited as a tentative hypothesis, then in the future I can only proceed hesitatingly and hypothetically, *conjecturing* that I probably ought to answer ‘68 + 57’ with ‘125’ rather than ‘5’. Obviously this is not an accurate account of the matter. There may be some facts about me to which my access is indirect, and about which I must form tentative hypotheses: but surely the fact as to what I mean by ‘plus’ is not one of them! To say that it is, is already to take a big step in the direction of skepticism. Remember that I **immediately and unhesitatingly** calculate ‘68 + 57’ as I do, and the meaning I assign to ‘+’ is supposed to *justify* this procedure. **I do not form tentative hypotheses**, wondering what I should do if one hypothesis or another were true. [p. 40]

The problem Kripke points to here is that facts about how I am disposed to apply a word (given that they are to be identified with brain states) not only fail to satisfy condition (2) of [J<sub>I</sub>], as the above objection maintains, but also fail to satisfy condition (1) of [J<sub>I</sub>]:

[J<sub>I</sub>] F justifies my applying word w thus-and-so, given what I mean by it, if and only if (1) I have *direct (non-inferential) access* to F, and (2) I can *deduce a priori* from F that I ought to apply w thus-and-so.

since I do not have *direct (non-inferential) access* to physical states of my brain.

Not only does [J<sub>I</sub>] rule out facts about my brain state from serving as meaning-constituting facts (since they are not directly knowable by me), but, more generally, any

fact that is not immediately accessible to me – such as facts about the grammar of our language that Chomsky credits us with *knowing implicitly*. So the argument is quite broad in its scope.

It is worth elaborating a bit on this point, as Chomsky's defense [1986, Chapter 4] of his theory as a straight solution (to the version of the skeptical problem applying not to words but to *sentences* of our language) has received a great deal of attention (and support) in the literature. Chomsky's theory purports to 'explain' linguistic competence in terms of the neurophysiological makeup of the brain. Mastery of a language or competence in its use, according to Chomsky, is 'a state of the individual mind/brain' [1986 p. 224] which consists in the 'internalization' of an assembly of 'rules of grammar' (a finite set of axioms which Chomsky conceives as codifying the potentially infinite character of language) which enable us to understand the meanings of indefinitely many sentences that we have never seen or heard before. These rules are part of an 'internal representational system', and are 'psychologically real' -- i.e., 'states of the mind/brain'. [1986 p. 221] 'They are factual in whatever sense statements about valence or chemical structure or visual equipment are said to be factual and involve truth claims.' [1986 p. 224] Chomsky contends that it is not part of an individual's mastery of his language to know these rules 'explicitly' or to be able to formulate them -- speakers barely recall what little grammar they learned at school. Rather, our knowledge of these rules is 'tacit'. [1986 p. 223]

But given that we do not have direct access to these rules, they fail to satisfy condition (1) in [J<sub>1</sub>], and so, the skeptic would conclude, they cannot serve as a straight solution. And even if we granted that (1) holds, that we have direct access to 'rules'

coming ‘before the mind’, the skeptic would not be silenced, for he would now resurrect his objection against candidate [A]. ‘To say that there is a general rule in my mind that tells me how to add in the future is only to throw the problem back on to other rules that also seem to be given only in terms of finitely many cases.’ [p. 22]<sup>22</sup>

#### *D. Primitive State*

It looks, at last, as if we have exhausted the realm of possible facts that can be taken to constitute meaning *addition* by ‘+’. So maybe, the skeptic hypothesizes, meaning *addition* by ‘+’ is ‘a **primitive state**, not to be assimilated with sensations or headaches or any “qualitative” state, nor to be assimilated to dispositions, but a state of a unique kind of its own’. [p. 51-53] Kripke takes this to be a final ‘desperate’ move one might make who agrees that the candidate facts considered and rejected exhaust the possibilities, yet refuses to accept the skeptical conclusion. [p. 51] ‘Surely’, such a person might protest, ‘meaning something by an expression has to be some kind of *state*. For how else can I be so confident that I do, at present, mean *addition* by “+”?’

The skeptic rejects the primitive state proposal for two reasons. The first has to do with the purported failure of such facts to satisfy clause (1) in [J<sub>1</sub>].

it leaves the nature of this postulated primitive state – the primitive state of “meaning addition by plus” – completely mysterious. It is **not supposed to be an**

---

<sup>22</sup> Even if we remove the justification constraint, the threat of a regress remains: What does our understanding of the alleged rules consist in? In having direct access to further rules? If so, what does our understanding of these further rules consist in? ... As we can see, we are off again on a regress, so that we are no better off with the alleged rules than we were without them; so the skeptic would conclude.

**introspectible state**, yet we are aware of it with some fair degree of certainty whenever it occurs. [p. 51]

The point here seems to be that once we abandon the idea that meaning so-and-so by w is some kind of ‘introspectible’ state – such as a qualitative state of consciousness -- it becomes a mystery how we can account for our *first-person knowledge* of meaning. How can we account for our confidence that we mean so-and-so by w if it is not an introspectible state? Kripke’s skeptic appears to be saying that we cannot explain our ability to *know directly* what we mean by ‘+’ if we endorse a *non-introspective* conception of the facts that determine what we mean by ‘+’.

The second reason the skeptic rejects the primitive state proposal has to do with its purported failure to satisfy condition (2) in [J<sub>1</sub>]. The objection runs: How can I deduce how I am to apply a word in an unlimited number of cases from a finite state of mine? How can such information be stored up or encoded in a finite state?

Even more important is the logical difficulty implicit in Wittgenstein’s skeptical argument. I think that Wittgenstein argues, not merely as we have seen hitherto, that introspection shows that the alleged ‘qualitative’ state of understanding is a chimera, but also that it is logically impossible (or at least that there is a considerable logical difficulty) for there to be a state of ‘meaning addition by “plus”’ at all. Such a state would have to be a finite object, contained in our finite minds. It does not consist in my explicitly thinking of each case of the addition table, nor even of my encoding each separate case in the brain. We lack the capacity for that. Yet (§195) “In a *queer* way” each such case already is “in some sense present”. ... What can that sense be? ... The proposal I am now discussing brushes such questions under the rug, since the nature of the “supposed” state is left mysterious. [p. 52]

The skeptic’s point here is that the primitive state proposal fails to answer the question: How could a primitive state, which would have to be a ‘finite object’ (contained in our

finite mind) be such as to reach out to a number of future uses of an expression and determine whether or not they are correct? How can a finite state of mind logically determine for a large (perhaps infinite) range of cases, how a term is to be applied? It doesn't look like that is possible.<sup>23</sup>

\*\*\*

This concludes Kripke's (or, more accurately, the skeptic's) reasons for rejecting [A] - [D], and hence, for concluding that there is no fact about me that corresponds to my meaning *addition* by '+'. Concerning the role of the justification constraint in reaching this conclusion, Kripke remarks:

there is a condition that any putative candidate for such a fact [that constitutes my meaning plus] must satisfy. It must, in some sense, **show how I am justified** in giving the answer '125' to '68+57'. The 'directions' mentioned in the previous paragraph, that determine what I should do in each instance, must somehow be 'contained' in any candidate for the fact as to what I meant. Otherwise, the sceptic has not been answered when he holds that my present response is arbitrary. **Exactly how this condition operates will become much clearer below**, after we discuss Wittgenstein's paradox on an intuitive level, **when we consider various philosophical theories as to what the fact that I meant plus might consist in**. There will be many specific objections to these theories. But all fail to give a candidate for a fact as to what I meant that would show that only '125', not '5', is the answer I 'ought' to give. [p. 11]

---

<sup>23</sup> Crispin Wright maintains that this set of problems (though he doesn't locate them in WRPL as I have here) forms one of Wittgenstein's central preoccupations in PI, and has yet to be given a satisfactory solution. In his words, the problem is: 'How is it possible to be effortlessly, non-inferentially and generally reliably authoritative about psychological states which have no distinctive occurrent phenomenology and which have to answer, after the fashion of dispositions, to what one says and does in situations so far unconsidered?'. [2001, p. 150]

I tried, in this section, to clarify the justification constraint's role in the skeptical argument – to show exactly how it is used to knock down all facts entertained in the skeptical argument. I took this task to be important because how it operates in the skeptical argument is not brought out as clearly as one could hope in WRPL, and it is crucial to get its role into focus before assessing the skeptical argument, both in its own right and as an interpretation of PI.

### 1.3 THE SKEPTICAL SOLUTION

In Chapter 3 of WRPL, Kripke outlines what he calls a '**skeptical solution**' to the skeptical problem. There are two components to the 'skeptical solution'. On the one hand, unlike a 'straight solution', the skeptical solution is not intended to refute the conclusion of the skeptical argument. What makes the position 'skeptical' is that it *accepts* the conclusion of the skeptical argument:

[SC] For any subject S, there is no fact about S that corresponds to his meaning *addition* by '+' (or anything by any word).

What makes it a 'solution' is that it nevertheless defends our ordinary *practice* of ascribing meaning and understanding to each other.<sup>24</sup> According to the skeptical solution, our ordinary practice of ascribing meanings to each other is not undermined by [SC], for, contrary appearances notwithstanding, it is not jeopardized by the failure of any

fact (state of an agent) to justify our applications of words in the way demanded by the skeptic:

[J<sub>I</sub>] F justifies my applying word w thus-and-so if and only if (1) I have *direct (non-inferential) access* to F, and (2) I can *deduce a priori* from F that I ought to apply w thus-and-so.

As noted earlier, [J<sub>I</sub>] is an **internalist** account of justification, which makes reference only to an individual. The skeptical solution insists that the failure of any fact to satisfy [J<sub>I</sub>] does not undermine our ordinary practice of meaning-attributions, because that practice can be justified in another, more plausible sense.

Any individual who claims to have mastered the concept of addition will be judged by the community to have done so if his particular responses agree with those of the community in enough cases, especially the elementary ones (and if his ‘wrong’ answers are not *bizarrely* wrong, as in ‘5’ for ‘68 + 57’, but seem to agree with ours in *procedure*, even when he makes a ‘computational mistake’). An individual who passes such tests is admitted into the community as an adder; an individual who passes such tests in enough other cases is admitted as a normal speaker of the language and member of the community. Those who deviate are corrected and told (usually as children) that they have not grasped the concept of addition. One who is an incorrigible deviant in enough respects simply cannot participate in the life of the community and in communication. [p. 91-92]

The more plausible sense in which our practice of meaning-attribution is justified, according to the skeptical solution, can be summed up as follows:

---

<sup>24</sup> I follow Scott Soames [1998 p. 313] in this interpretation of the terms ‘skeptical’ and ‘solution’.

[J<sub>E</sub>] ‘S means M by w’ is warranted/justified iff S’s applications of w accord with those of his linguistic community, and M is the meaning w has in S’s linguistic community.

On this **externalist** (social) account of justification, **justification (or assertibility)**<sup>25</sup> **conditions** for meaning-ascriptions involving ‘+’ -- i.e., conditions under which it is licensed to make such assertions as ‘S means *addition* by “+”’ -- are given as follows: S’s linguistic community is warranted/justified in asserting of S that he means *addition* by ‘+’ if and only if S has satisfactorily often enough come up with the answer that most of the rest of his fellow speakers give to addition queries, and they mean *addition* by ‘+’.<sup>26</sup>

[J<sub>E</sub>] is an externalist account of justification because, unlike [J<sub>I</sub>], which grounds justification-conditions for a subject’s applying a word in facts *internal* to the subject, [J<sub>E</sub>] grounds justification-conditions for a subject’s applying a word in facts *external* to the subject – specifically, in communal practices involving that word. Here, attributions of meaning call on no special state of a subject’s mind, but are licensed by social practices. [J<sub>E</sub>] is intended to capture the intuitive idea that the meaning a subject attaches to a word might not be the meaning that it has in his common language; only in so far as it is are we entitled to say that he understands the word. His applications must accord

---

<sup>25</sup> If we speak properly, Kripke notes on p. 73, ‘we should not speak of conditions of “assertion”, but rather, more generally, of the conditions when a move (a form of linguistic expression) is to be made in the “language game”. If, however, we allow ourselves to adopt an oversimplified terminology more appropriate to a special range of cases, we can say that Wittgenstein proposes a picture of language based, not on truth-conditions, but on *assertibility conditions* or *justification conditions*: under what circumstances are we allowed to make a given assertion?’

<sup>26</sup> There is ‘no objective fact that explains our agreement’, Kripke adds. [p. 97] For Wittgenstein, an “explanation” of this kind ignores his treatment of the sceptical paradox and its solution ... our license to say of each other that we mean addition by “+” is part of a “language game” that sustains itself only because of the brute fact that we generally agree.’ [p. 97]

with the appropriate *regularities* observed among other members of his linguistic community.

Of course, this interpretation presupposes sufficient communal agreement on how the word is applied. In absence of such agreement, the word would be devoid of meaning, and we would not be entitled to ascribe mastery of it to a subject. But where the required agreement exists, it supplies the justification conditions for attributing meaning to a subject, provided he exemplifies sufficiently non-deviant behavior. [WRPL p. 95]

There are at least two delicate issues regarding the skeptical solution that stand in need of clarification. These will form the topic of discussion in this section. One issue concerns the very coherence of the skeptical conclusion [SC]. On the face of it, [SC] appears to be *self-undermining*: if there is no fact about S that his meaning *addition* by ‘+’ (or anything by any word) corresponds to, for any subject S, then how can anyone succeed in stating the skeptical conclusion without contradicting himself?

It all depends on how we interpret the skeptical conclusion. A number of interpretations have been offered in the literature; among them some of the following:

- [a] (Radical skepticism) All words are meaningless.
- [b] (Error theory) All meaning-ascriptions are false.
- [c] (Expressivism) Truth/falsity and factuality/non-factuality is not properly predicable of meaning-ascriptions.
- [d] (Anti-individualism) Meaning-ascriptions cannot be analyzed in terms of facts about the subject *alone*.

- [e] (Anti-reductivism) Meaning-ascriptions cannot be analyzed in terms of any collection of natural, non-semantic facts.

In the next section (§1.3.1), I will flesh out each of these interpretations. Among them, only [a] - [c] are self-defeating or incoherent -- we will look at some arguments that bring this out – and, though [d] and [e] are coherent, I will show that, along with the other candidates, they lack crucial features that Kripke attributes to the skeptical conclusion.

Though Kripke's formulation of the skeptical conclusion has a confusing tendency to vary, it should be apparent, from my discussion of the skeptical argument in the preceding section, what I take it to be (i.e., what I interpret [SC] to mean):

- [f] There is no state of an agent, analyzable or non-analyzable/primitive, that satisfies [J<sub>1</sub>].

The second interpretive issue that stands in need of clarification concerns the relation between the skeptical solution and Kripke's interpretation of Wittgenstein's 'private-language argument'. Kripke sees the conclusion of the private-language argument as a 'corollary' [p. 68] to the skeptical solution. But in what sense is it a 'corollary'? In §1.3.2, I flesh out the connection.

### **1.3.1 Interpreting the Skeptical Conclusion**

### *a. Radical Skepticism*

Some commentators<sup>27</sup> have interpreted the skeptical conclusion [SC] to be the *radical* skeptical thesis that:

[a] All words are meaningless.

Though [a] is not *internally* self-defeating (inconsistent), it is *pragmatically* self-defeating. If one has succeeded in stating (asserting) [a], believing it, or arguing for it, then one has *refuted* it.<sup>28</sup>

Can such thoroughgoing skepticism about meaning be the conclusion of the skeptical argument? Kripke himself addresses this question:

.... Wittgenstein holds, with the skeptic, that there is no fact as to whether I mean plus or quus. But if this is to be conceded to the skeptic, is this not the end of the matter? What *can* be said on behalf of our ordinary attributions of meaningful language to ourselves and to others? Has not the **incredible and self-defeating conclusion** that all language is meaningless, already been drawn? [p. 71]

In the ensuing discussion, Kripke maintains that *no such conclusion follows*, because Wittgenstein merely wishes to deny the existence of facts ‘in the world’ [p. 72] that

---

<sup>27</sup> See, e.g., Rupert Read [2000] and George Wilson [1998, 1994].

<sup>28</sup> As Scott Soames points out. Thus, he writes:

The skeptic’s position is self undermining not in the sense that it is internally inconsistent, but in the sense that it is inconsistent with the supplementary claim (which must be rejected by the skeptic) that he, or anyone else, has stated the skeptical position, argued for it, or believed the conclusion. (Of course, since rejecting that so and so is also a propositional attitude, crucially involving the notion of content, the skeptic must also reject – as we would put it – the claim that he has rejected anything.) [1998 p. 317]

philosophers misleadingly associate with meaning-attributions, not the propriety of the meaning-attributions themselves.

All that is needed to legitimize assertions that someone means something is that there be roughly specifiable circumstances under which they are legitimately assertable, and that the game of asserting them under such conditions has a role in our lives. **No supposition that ‘facts correspond’ to those assertions is needed.** [p. 77]

This point is also emphasized in an earlier passage:

I have said that Wittgenstein’s solution to his problem is a skeptical one. He does not give a ‘straight’ solution, pointing out to the silly skeptic a hidden fact he overlooked, a condition in the world which constitutes my meaning addition by ‘plus’. In fact, **he agrees with his own hypothetical skeptic that there is no such fact, no such condition in either the ‘internal’ or ‘external’ world.** Admittedly, I am expressing Wittgenstein’s view more straightforwardly than he would ordinarily allow himself to do. For in denying that there is any such fact, might we not be expressing a philosophical thesis that doubts or denies something everyone admits? We do not wish to doubt or deny that when people speak of themselves and others as meaning something by their words, as following rules, they do so with perfect right. We do not even wish to deny the propriety of an ordinary use of the phrase ‘the fact that Jones meant addition by such-and-such a symbol’, and indeed such expressions do have ordinary perfect uses. **We merely wish to deny the existence of the ‘superlative fact’ that philosophers misleadingly attach to such ordinary forms of words, not the propriety of the forms of words themselves.** [p. 69]

We need to distinguish two characters in these passages: Kripke’s Wittgenstein (KW) and Kripke’s skeptic. The *skeptic* plays the role of *Wittgenstein’s interlocutor* in PI. Kripke is saying that, though Wittgenstein agrees with the hypothetical skeptic that there is no ‘superlative fact’ [p. 66, 69] that corresponds to meaning anything by any word, this is not the same as saying that he agrees with [a]. What I take, more specifically, to be the claim KW agrees with the skeptic on is:

[f] There is no state of an agent, analyzable or non-analyzable, that satisfies [J<sub>1</sub>].

By ‘superlative fact’ in the above passage, I take Kripke to mean a fact that ‘contains’ a criterion of correctness; more specifically, a fact that satisfies [J<sub>1</sub>]. Though KW accepts [f], he does not, unlike the skeptic, insist that [J<sub>1</sub>] be met: failure to satisfy [J<sub>1</sub>], as we have seen, is the skeptic’s ultimate reason for drawing [SC]. Rather, according to Kripke:

[Wittgenstein] claims (in §§183-93), the appearance that our ordinary concept of meaning demands such a fact is based on a philosophical misconstrual – albeit a natural one – of such ordinary expressions as ‘he meant such-and-such’, ‘the steps are determined by the formula’, and the like. [p. 65-6]

Kripke does not detail the nature of this ‘philosophical misconstrual’. This is an issue I will return to in Chapter 4, where I fill in the missing details.

### ***b. Error Theory***

An *error theory* about a fragment of a particular discourse, consisting of *declarative* sentences of that discourse, takes the fragment’s semantical appearances at face value: declarative sentences *describe objective facts* and *express representational thoughts*. However, according to the error theorist, because nothing actually corresponds to the

facts so described, all the fragment's declarative sentences are *false*: they are founded upon metaphysical error.

Such a view is adopted by, for instance, John Mackie about ethics.<sup>29</sup> According to Mackie, the declarative sentences of moral discourse purport to describe a range of objective prescriptions: constraints on action that are somehow built into the fabric of nature. Hence, they *presuppose* the existence of moral values, with something like the objectivity of material objects dealt with by physical science. But it can be shown (Mackie held) that there are no such things. So an ordinary user of moral language is typically making claims that involve a mistake about what the world contains. Mackie sees moral thought and language much as an atheist sees religious talk and language. The believers are not *conceptually* confused (they are not making a *semantic* mistake); rather, they are *ontologically* mistaken (the error is one of '*metaphysical superstition*'<sup>30</sup>).

Error theories have surfaced, not only in ethics, but across a wide variety of discourses.<sup>31</sup> It might be thought that the goal of the skeptical argument is to establish an error theory for the discourse of meaning, i.e., to show that:

[c] All meaning-ascriptions are *false*.

---

<sup>29</sup> See *Ethics: Inventing Right and Wrong*, Harmondsworth: Penguin, 1977.

<sup>30</sup> Wright's expression [1992 p. 5].

<sup>31</sup> They can be found in the treatment of color by Cartesian philosophers such as Malebranche and Arnauld, and more recently in the work of Paul Boghossian and David Velleman. According to these philosophers, colors are found only in the mind; the world itself is not really colored, but we naturally treat it as if it is. Paul Churchland defends an error theory about psychological talk concerning the mind. See his 'Eliminative Materialism and the Propositional Attitudes', *Journal of Philosophy*, 1981. Error theories have also surfaced in the philosophy of math, philosophy of science, and in aesthetics.

But this cannot be the correct interpretation of [SC]. Aside from the fact that there is no textual evidence for this reading, an error theory of meaning-attributions, in contrast with other regions of discourse, is self-defeating, as the argument below shows<sup>32</sup>

- i. For any S and M: ‘S means M by w’ is false. (error theory of meaning discourse)
- ii. So, ‘S does not mean M by w’ is true.
- iii. So, w has no meaning, i.e., all words are meaningless.

But, as we saw in the last section, the skeptical conclusion is not supposed to be so radical as to be self-defeating. So the skeptical conclusion cannot be [b].

---

<sup>32</sup> Paul Boghossian [1989 p. 523] and [1990 p. 174-5] offers the following argument to show that an error theory of meaning is incoherent (my summary).

- i. For any S and p: ‘S means that p’ is false. (error theory of meaning discourse)
- ii. A sentence of the form ‘S means that p’ is true iff S means that p. (This follows from the truth predicate’s disquotational properties.)
- iii. For any S: ‘S’ has no meaning, i.e. no sentence is meaningful. (This follows from (i) and (ii))
- iv. But what (i) says is that some sentences -- namely meaning-attributing sentences -- are false, and a sentence cannot be false unless it is meaningful to begin with.
- v. Hence (i) is incoherent, since it entails a contradiction.

Note that the argument assumes we are employing our *ordinary notions of truth and falsity*, according to which to say of a sentence s that s is true (or false) is to assume it is meaningful. But, as Scott Soames [1998 p. 344 n. 8] points out:

... the radical skeptic is free to replace the ordinary notions of truth and falsity with Tarski-type substitutes for his own language that do not presuppose any intentional notions. Let ‘TRUE’ and ‘FALSE’ be such predicates. A skeptic who says that there is a TRUE sentence written on the blackboard says something trivially equivalent to the following: either ‘Snow is white’ is written on the blackboard and snow is white, or ‘Grass is green’ is written on the blackboard and grass is green, etc. (Sentences of other languages could not be dealt with using the same predicate.) When this notion of truth is employed the skeptic will accept the inference from s to the claim that s is TRUE, and vice versa. He will characterize meaning-ascriptions as not TRUE and, depending on how FALSITY is defined, he may even characterize them as FALSE. However, now there is no contradiction in claiming that a sentence is both meaningless and TRUE (or FALSE). Similarly, the claim that there are conditions under which a sentence is TRUE, as well as the claim that a sentence has Truth-conditions, are not conceptually tied to any claim about the meaning of the sentence.

### *c. Expressivism*

*Expressivism* (or, as it is sometimes called, *emotivism*, *noncognitivism*, or *projectivism*) contrasts with an error theory in contesting what an error theory grants: that the declarative sentences of a disputed discourse *describe objective facts* or *express representational thoughts*. Rather than interpreting utterances in the contested discourse as describing facts, the expressivist urges us to interpret them as executing a variety of non-descriptive tasks, such as expressing emotions, attitudes, commands, or prescriptions. For the expressivist, superficial ‘grammatical’ similarities (i.e., similarities in syntactic form) of certain types of utterance *mislead us* into thinking that it is their business to describe facts. But, as closer scrutiny of their *role* in language will reveal, they have an entirely different function, and should not be assimilated to statements whose function it is to describe facts.

Expressivist theories have been advanced mainly in moral philosophy.<sup>33</sup> Along with the error theorist, the moral expressivist denies that there are objective moral facts which are the source of the truth-values of moral judgments (‘X is morally acceptable’, ‘X is right’, ‘X is good’, ‘X ought to be done’, etc.), as snow’s being white (state of affairs in the world) is the source of the truth value of the judgment ‘Snow is white’. But, unlike the error theorist, who infers from this that moral ‘judgments’ are *false*, the expressivist infers they are not descriptive statements and do not express beliefs, and hence that

truth/falsity is *not* properly predicable of them. Moral ‘judgments’, he maintains, serve to prescribe or caution against certain forms of conduct, and thereby to express some *non-cognitive* mental state of the judge -- a moral *attitude*. Beliefs and other representational thoughts purport to *represent* the world as being a certain way, and so are thought to have a *truthmaker* (state of affairs in the world that makes the belief true). Saying that X is right or wrong, good or bad, on the other hand, according to the expressivist, is supposed to be like saying ‘Hooray for X!’ or ‘Boo for X!’. These statements do not ascribe a genuine property to X (and hence do not have a truthmaker), as the statement ‘X is good’ is thought to (in virtue of its surface structure).<sup>34</sup>

Expressivism *about meaning-attributions* is the view that meaning-attributions do not *describe* states of affairs. Though they look like descriptive claims, being declarative sentences, they really have an entirely different role/function. In the case of ‘+’, it might be one of making a *prediction* about how a subject will use ‘+’. Or it might be one of *prescribing* conduct: ‘S should apply “+” in such and such a way’. Or,<sup>35</sup> it might be one of *commending* S: taking him into one’s linguistic community, certifying him as a competent user of ‘+’, and licensing him to use ‘+’ to do what we call ‘adding’.

---

<sup>33</sup> Such a view is adopted by, for instance, A. J. Ayer and R. M. Hare. See A. J. Ayer, *Language, Truth and Logic*, Harmondsworth, Middlesex, 1971, Ch. 6, and R. M. Hare, *The Language of Morals*, Oxford: Oxford University Press, 1952.

<sup>34</sup> Expressivist theories have surfaced in a wide range of discourses. For example, in the philosophy of mathematics, they have surfaced in work of the later Wittgenstein, and in David Hilbert’s *formalism* (the view that mathematics is a game with meaningless marks, manipulated according to formal rules). They have also taken the form of *instrumentalism* in the philosophy of science (according to which theoretical statements function as tools for deriving predictions about future experience). And they have surfaced in the philosophy of mind, in the work of Daniel Dennet, according to whom judging that someone has intentional states is taking the ‘intentional stance’ toward that person. See *The Intentional Stance*, Cambridge, MA, 1987.

<sup>35</sup> As pointed out by Scott Soames [1998 p. 322].

Many leading commentators on WRPL<sup>36</sup> have interpreted [SC] along expressivist lines, as follows:

[c] Truth/falsity (or factuality/non-factuality) is not properly predicable of meaning-ascriptions.

On this interpretation, the moral or the skeptical argument is that our hunt for the fugitive meaning-constituting fact is entirely misconceived because meaning-ascriptions (e.g. ‘Jones means *addition* by “+”’) do not make any claims about the world.

[c], however, cannot be the correct interpretation of [SC]. To begin with, a number of arguments have been offered in the literature (e.g., Scott Soames [1998], Cripsin Wright [1984] and Paul Boghossian [1989]) that show [c] is incoherent, and, as we saw earlier, [SC] is not supposed to be so radical as to be incoherent.

Here, e.g., is Soames’ argument:

... the non-descriptive version of the skeptical solution is incoherent. One way to bring out this incoherence is with the following argument: Consider the sentence, ‘no meaning ascription expresses a proposition’, used in stating this version of the skeptical solution. It makes a theoretical claim about a certain class of sentences, and so must itself be proposition-expressing and fact-stating. Thus, its negation, ‘some meaning ascription expresses a proposition’, ought to have the same status. This sentence is an existential generalization whose instances are sentences of the form ‘s is a meaning ascription and s expresses a proposition’. Surely if the existential generalization expresses a proposition, then some of its instances do too, in which case some instances of the form ‘s expresses a proposition’ must express propositions. But these are meaning-ascriptions of a certain sort, which, by the non-descriptive version of the skeptical solution, do not express

---

<sup>36</sup> See, for example, Boghossian [1989 p. 524-527], Wright [1984 p. 761], and McGinn [1984 p. 64]. For discussions of why this interpretation cannot be correct, see Paul Horwich [1998, Chapter 10], Scott Soames [1998], and Alex Byrne [1992].

propositions. Thus we have it that some sentences both express propositions and do not express propositions. [1998 p. 324]

Crispin Wright's strategy (which is also adopted by Boghossian) is to show that meaning nonfactualism entails *global* nonfactualism, and hence is implausible, i.e. that (1) entails (2):

- (1) For any S and p, "S" means that p' is non-factual.
- (2) For any S, 'S' is non-factual.

Here is Wright's argument in his [1984] and [1992] papers:

The elegance of Kripke's interpretation [that meaning-ascriptions have no truth-conditions] does not long conceal its difficulties. One immediate difficulty is presented by the meaning truth platitude. If the truth value of S is determined by its meaning and the state of the world in relevant respects, then nonfactuality in one of the determinants can be expected to induce nonfactuality in the outcome. ... A projectivist view of meaning is thus, it appears, going to enjoin a projectivist view of what it is for a statement to be true. Whence, unless it is, mysteriously, possible for a projective statement to sustain a biconditional with a genuinely factual statement, the disquotational schema 'P' is true iff P' will churn out the result that all statements are projective. [1984 p. 769]

... the irrealism about meanings, rules and their requirements imposed by the skeptical paradox -- the thesis that there are 'no facts of the matter' as far as rules and meanings are concerned -- must necessarily inflate, it may seem, into *global* irrealism: the thesis that there are no facts of the matter *anywhere*. It is the merest truism that whether any sentence is true is a function, in part, of what it says. If there are no substantial facts about what sentences say, there are no substantial facts about whether or not they are true. Thus, irrealism about meaning must enjoin an irrealism about truth, wherever the notion is applied. And irrealism about truth, wherever the notion is applied, is irrealism about all assertoric discourse. [1992 p. 211]<sup>37</sup>

---

<sup>37</sup> Wright's argument can be summarized as follows:

- i Suppose for any S and p: "S" means that p' is non-factual. (meaning-nonfactualism)

There is, moreover, ample textual evidence that Kripke did not intend the skeptical conclusion to be interpreted as [c]. In addition to the evidence cited above (quotes from WRPL, pp. 65-66, 69 and 77), there is the following remark:

Wittgenstein's skeptical solution concedes to the skeptic that no 'truth-conditions' or 'corresponding facts' in the world exist that make a statement like 'Jones, like many of us, means addition by "+" true. Rather we should look at how such assertions are *used*. Can this be adequate? Do we not call assertions like the one just quoted 'true' or 'false'? Can we not with propriety precede such assertions with 'It is a fact that' or 'It is not a fact that'? Wittgenstein's way with such objections is short. **Like many others, Wittgenstein accepts the 'redundancy' theory of truth: to affirm that a statement is true (or, presumably, to precede**

- 
- ii Then for any S: "'S" is true' is non-factual. (Since the truth-condition of any sentence is (in part) a function of its meaning, nonfactualism about meaning entails nonfactualism about truth.)
  - iii So for any S, 'S' is non-factual. (This follows from the disquotational scheme, together with the assumption that if the sentence A is non-factual, and 'A iff B', then the sentence B is non-factual.)
  - iv So meaning- nonfactualism entails *global* nonfactualism.

Paul Boghossian takes this argument a step further [1989 p. 525ff and 1990 p. 175ff], and tries to show that meaning nonfactualism is *incoherent*. The incoherence, he maintains, is generated from the clash between what you have to suppose about truth in order to frame a non-factualist thesis about *anything* and what you have to suppose about truth as a result of accepting non-factualism about *meaning*. Below is my summary of his argument.

- i Suppose that for any S and p: "'S" means that p' is non-truth-conditional. (meaning-nonfactualism)
- ii Then for any S and p: "'S" has the truth-condition that p' is non-truth-conditional. (Since the truth condition of any sentence is (in part) a function of its meaning, nonfactualism about meaning will enjoin nonfactualism about truth-conditions.)
- iii So, judgments about a sentence's truth value *cannot be factual*.
- iv On the *deflationary* understanding of truth, for a sentence to possess truth-conditions, it is both necessary and sufficient that it be *significant* and *declarative* in form.
- v But *non-factualism* denies of some significant declarative sentence ("S" means that p') that it is truth-conditional, so the concept of truth in terms of which a non-factualist thesis is framed can't be the deflationary concept. Non-factualism presupposes a concept of truth richer than the deflationary: it is committed to holding that the predicate 'true' stands for some language-independent property, i.e. that truth is 'robust'.
- vi If truth is robust, then judgments about a sentence's truth value *are factual*.
- vii (vi) contradicts (iii). Therefore non-factualism about meaning is incoherent.

For a critique of Boghossian's extension of Wright's argument, see Soames [1998 p. 251ff] and Wright [1992, Appendix].

it with ‘It is a fact that...’) is simply to affirm the statement itself, and to say it is not true is simply to deny it: (‘p’ is true = p). ... We call something a proposition, and hence true or false, when in our language we apply the calculus of truth functions to it. That is, it is just a primitive part of our language game, not susceptible of deeper explanation, that truth functions are applied to certain sentences. [p. 86]

As noted earlier, we need to distinguish two characters in these passages: Kripke’s Wittgenstein and Kripke’s skeptic. The *skeptic* plays the role of *Wittgenstein’s interlocutor* in PI. Neither Kripke’s Wittgenstein (proponent of the skeptical solution) nor Kripke’s skeptic (proponent of the skeptical argument) is rejecting the propriety of preceding meaning-ascriptions with ‘It is true that’ or ‘It is a fact that’, i.e., both grant that one can correctly say things like ‘*It is a fact* that Jones means addition by “+” or ‘*It is true* that Jones means addition by “+”’. No evidence has been offered to the contrary. In other words, both Kripke’s skeptic and KW endorse what, in contemporary terminology, is known as **minimalism**<sup>38</sup> about truth. Minimalism differs from an error or expressivist theory by denying that (i) and (ii) follows from the stated hypothetical:

If there are no objective facts which are the source of the truth-values of statements in a particular discourse, then:

- (i) (Error theory) Those statements are false.
- (ii) (Expressivism) Truth/falsity (or factuality/non-factuality) is not properly predicable of them.

---

<sup>38</sup> For a discussion and defense of this theory of truth, see Paul Horwich: *Truth* (Oxford: Basil Blackwell, 1990), and Crispin Wright: *Truth and Objectivity* (Cambridge Mass: Harvard University Press, 1992). Simon Blackburn’s ‘quasi-realism’ is also a form of minimalism. See his *Essays on Quasi-Realism*, New York, 1993 and *Ruling Passions*, Oxford, 2000.

- (iii) (Minimalism) Truth/falsity (or factuality/non-factuality) may be properly predicable of them (pace (ii)), and they need not be false (pace (i)).

According to the minimalist, truth/falsity (or factuality/non-factuality) is properly predicable of a statement if it is governed by agreed standards of warrant, and is embeddable within conditionals, disjunctions, conjunctions, and other connectives, and within contexts of propositional-attitude verbs. Since meaning-ascriptions satisfy these conditions, truth/falsity (or factuality/non-factuality) is properly predicable of them. On this view, there is nothing incoherent about saying '*It is true* that Jones means addition by "+"', or '*It is a fact* that Jones means addition by "+"', even if there are no objective facts that are the source of the truth-values of these statements. These are permissible moves in our language game. We simply must beware that in making these statements, we are not investing them with ontological weight / the metaphysical commitments associated with realism.

This is where Kripke's skeptic differs from KW. Kripke's skeptic, unlike Kripke's Wittgenstein, mistakenly inflates these statements with a false ontology. He assumes that they commit us to a 'superlative fact': a state that satisfies [J]. After searching in vain for it, he concludes that no such fact exists. The skeptical solution is supposed to show that this result is not as radical or paradoxical as it might seem. It draws us back from the abyss, by showing that no such fact is needed to legitimate our ordinary practice of meaning-attribution. To this end, it makes essential use of the notion of a 'community'. Jones' linguistic community is justified in asserting that '(It is true that) Jones means *addition* by "+"', not because they have evidence of the mythical fact, but because Jones'

behavior involving ‘+’ matches theirs (and they mean *addition* by ‘+’). That is the way the language game of attributing meanings works. The skeptic (Wittgenstein’s interlocutor) was seduced, as usual, by misleading philosophical pictures; but once we see them for what they are, by examining the role and function of attributions of meaning in our language, we can rest assured that it is a ‘fact’ that agents mean what they do by their words, though not the ‘superlative’ one that our false pictures, quite naturally, led us to think existed.

#### *d. Anti-individualism*

On the interpretation of [SC] that I have offered, the ‘superlative fact’ whose existence is being denied by [SC] is one that satisfies [J<sub>1</sub>]. According to other interpreters, by denying the existence of a ‘superlative fact’, the skeptic does not mean to be denying this, but rather that:

[d]     Meaning-ascriptions cannot be analyzed in terms of facts about the subject  
*alone.*

This position (known as **anti-individualism** in the literature) has become familiar through the work of Hilary Putnam and Tyler Burge.<sup>39</sup> It implies that if Jones means *addition* by ‘+’ then there is no fact *solely about Jones* in virtue of which this is the case.

---

<sup>39</sup> See Putnam [1973] and Burge [1979].

This leaves open the possibility that what a subject means by a word *w* is analyzable in terms of a *relation* between him and his linguistic community or other facts external to him. The relation might, e.g., be one of agreement in use of *w*. In this case, the fact that constitutes Jones' meaning *addition* by '+' would be partly a fact about Jones and partly a fact about his linguistic community.

Unlike our earlier interpretations of [SC], [d] is not self-defeating or otherwise implausible. Also, unlike our earlier interpretations, one can find some textual support for this interpretation. To begin with, the skeptical problem is always presented in *individualistic* terms: Is there any fact *about me* that constitutes my meaning *addition* by '+'?

Wittgenstein's challenge can be presented to me as a question about *myself*: was there some past fact about **me** – what I meant by 'plus' – that mandates what I should do now? [p. 15]

A negative answer to this question (as given in the passages below) is *compatible* with anti-individualism.

The skeptic holds that no fact about **my** past history – nothing that was ever in my mind, or in my external behavior – establishes that I meant plus rather than quus. [p. 13]

The skeptical challenge ... purports to show that nothing in **my** mental history or past behavior – not even what an omniscient God would know – could establish whether I meant plus or quus. But then it appears to follow that there is no *fact* about **me** that constituted my having meant plus rather than quus. [p. 21]

Wittgenstein's skeptic argues that he knows of no fact about an **individual** that could constitute his state of meaning plus rather than quus. [p. 39]

It turns out that the skeptical solution does not allow us to speak of a single individual, considered **by himself and in isolation**, as ever meaning anything. [p. 69]

Moreover, [d] finds support in Kripke's discussion of the relationship between the skeptical solution and the private-language argument.

Following §243, a private-language is usually defined as a language that would be logically impossible for anyone else to understand. The private-language argument is taken to argue against the possibility of a private-language in this sense. This conception is not in error, but it seems to me that the emphasis is somewhat misplaced. What is really denied is what might be called the 'private model' of rule-following, that the notion of a person following a given rule is to be analyzed simply in terms of facts **about the rule follower and the rule follower alone, without reference to his membership in a wider community.**

Kripke's interpretation of the private-language argument, it appears from this passage, is that meaning and rule-following cannot be analyzed in terms of facts *about an individual alone*. This does not exclude the possibility that meaning and rule-following is analyzable in terms of facts about an individual *and his relation to his linguistic community*.

And finally, [d] finds support in Kripke's comparison of Wittgenstein's 'skeptical solution' to Hume's about causation. According to Kripke, Wittgenstein's argument against private-language has a structure similar to Hume's against private-causation.

Naively, one might suppose that whether a particular event a causes another particular event b is an issue solely involving the events a and b alone (and their relations), and involves no other events. If Hume is right, this is not so. Even if God were to look at the events, he would discern nothing relating them other than that one succeeds the other. Only when the particular events a and b are thought of as subsumed under two respective event types, A and B, which are related by a generalization that *all* events of type A are followed by events of type B, can a be

said to ‘cause’ b. **When the events a and b are considered by themselves alone, no causal notions are applicable.** This Humean conclusion might be called: the impossibility of private causation. [pp. 67-68]

After the skeptical solution has been seen to be unanswerable on its own terms, a skeptical solution is offered, containing all we can salvage of the notion of causation. **It is just a feature of this analysis that causation makes no sense when applied to two isolated events,** with the rest of the universe removed. Only inasmuch as these events are thought of as instances of event types related by a regularity can they be thought of as causally connected. If two particular events were somehow so *sui generis* that it was logically excluded that they be placed under any (plausibly natural) event types, casual notions would not be applicable to them. [p. 68]<sup>40</sup>

Despite what all this evidence appears to suggest, however, I don’t think Kripke intended the skeptical conclusion to be interpreted as [d].<sup>41</sup> The following passage makes this apparent [p. 111].

**Wittgenstein’s theory should not be confused with a theory that, for any *m* and *n*, the value of the function we mean by ‘plus’, is (by definition) the value that (nearly) all the linguistic community would give as the answer. Such a theory would be a theory of the *truth-conditions* of such assertions as ‘by “plus” we mean such-and-such a function, which when applied to 68 and 57 as arguments, yields 125 as value.’ ... The theory would assert that 125 is the value of the function meant for given arguments iff ‘125’ is the response that nearly everyone would give, given these arguments. Thus the theory would be a social, or community-wide version of the dispositional theory, and would be open**

---

<sup>40</sup> The following passage has also been cited as evidence for [d]:

[Wittgenstein] does not give a ‘straight’ solution, pointing out to the silly skeptic a hidden fact he overlooked, a condition in the world which constitutes my meaning *addition* by ‘+’. In fact, he agrees with his own hypothetical skeptic that there is no such fact, no such condition in either the ‘internal’ or ‘**external**’ world. [p. 69]

One might interpret ‘external world’ as a reference to my linguistic community or other factors external to me, but I think it is actually a reference to my *external behavior*, and so cannot be used to support [d]. Kripke says on p. 13 that ‘The skeptic holds that no fact about my past history – nothing that was ever in my mind, or in my **external behavior** – establishes that I meant plus rather than quus.’ And again on page 21 he says that ‘nothing in my own mental history or **external behavior** will answer the sceptic who supposes that in fact I meant quus’. I think he is simply remaking this point on p. 69.

<sup>41</sup> Paul Horwich [1998 ch. 10] and Scott Soames’ [1998a] also discuss and reject this as a plausible reading of the skeptical conclusion.

to at least some of the same criticisms as the original form. I take Wittgenstein to deny that he holds such a view, for example, in RFM V, §33 [VII, §40]: “Does this mean, e.g., that the definition of the same would be this: same is what all or most human beings ... take for the same? – Of course not.” (See also PI, p. 226, “Certainly the propositions, “Human beings believe that twice two is four” and “Twice two is four” do not mean the same”; and see also §§240-1.) One must bear firmly in mind that Wittgenstein has no theory of truth-conditions -- necessary and sufficient conditions -- for the correctness of one response rather than another to a new addition problem. Rather he simply points out that each of us *automatically* calculates new addition problems (without feeling the need to check with the community whether our procedure is proper); that the community feels entitled to correct a deviant calculation; that in practice such deviance is rare, and so on. Wittgenstein thinks that these observations are enough to illuminate the role and utility in our lives of assertion about meaning and determination of new answers. **What follows from these assertibility conditions is not that the answer everyone gives to an addition problem is, by definition, the correct one**, but rather the platitude that, if everyone agrees upon a certain answer, then no one will feel justified in calling the answer wrong. [p. 111]

The claim being made here is that the considerations adduced by the skeptic to show that facts about how I am disposed to apply a term don't determine meaning carry over to social facts of the same type. Take for example the finitude objection. The practice of the community will cover only a *finite* portion of possible applications of addition, and so, as in the individual case, cannot suffice to determine which rule is being followed, on a communal level. So how can communal practices fix the meaning of 'plus', or indeed of any term? A community is made up of individuals, so how is it determined what the community as a whole means by an expression, if the practice of each of the individuals who make it up is finite and therefore indeterminate? (The error objection, Kripke maintains, can be likewise generalized.)

It follows, according to Kripke, that agreement with communal practices cannot be what makes a given meaning-ascription *true*; rather, it is what *justifies* it. If Jones' practice (on the finite number of occasions on which it will have been displayed) accords

with the linguistic practice of his community (on the finite number of occasions on which it will have been displayed), then his linguistic community is *justified* (defeasibly entitled) in asserting that Jones means *addition* by '+', assuming they do.

If agreement with communal practices could serve as the truth conditions for meaning-ascriptions, then it would follow that the extension of 'plus' would be determined by communal verdicts:

- (\*) For any ordered triple of numbers  $\langle p, q, r \rangle$ ,  $\langle p, q, r \rangle$  is in the extension of 'plus' iff our linguistic community would judge that  $\langle p, q, r \rangle$  satisfies 'p+q=r'.

But (\*), Kripke is saying in the passage, cannot be correct: the finitude objection shows that the left-to-right direction of the biconditional fails, and according to the error objection, the linguistic community might be prone to error and wrongly judge of some  $\langle p, q, r \rangle$  that  $p+q=r$ . (\*), on the other hand, makes community wide arithmetic error impossible since communal practices define the extension of '+'. So we can conclude by *reductio* that (\*) is incorrect. As Kripke puts the point in the passage: 'What follows from these assertibility conditions is not that the answer everyone gives to an addition problem is, by definition, the correct one, but rather the platitude that, if everyone agrees upon a certain answer, then no one will feel justified in calling the answer wrong.'<sup>42</sup>

---

<sup>42</sup> George Wilson maintains that, on its most naturalistic interpretation, the skeptical conclusion is *incompatible* with the perspective of the skeptical solution, for, (i) implicit in the skeptical solution is the idea that ascriptions of meaning about a given speaker with proper warrant are directly committed to the claim that the speaker's relevant linguistic dispositions are properly aligned with relevant standards of correctness settled upon within the linguistic community; and (ii) given that meaning ascriptions carry this commitment, according to the positive strands of the skeptical solution, they cannot be coherently regarded there as having non-factual content. They state or imply purported facts about the relations between the

Whether this claim – that agreement with communal practices fails to serve as the truth condition for meaning-ascriptions – is correct is an issue I will return to in Chapter 2. There, I will defend the view that, with one exception involving a very small class of cases, it is incorrect, so long as we introduce a *ceteris paribus* clause into (\*).

There is another point that the skeptic has at his disposal against [d]: I do not have *direct (non-inferential) access* to how my linguistic community is disposed to use words, nor can I deduce *apriori* from such facts (given that I must observe how my applications relate to communal applications) that I ought to apply a word thus-and-so. So such facts fail to satisfy [J<sub>I</sub>], and any facts that fail to satisfy [J<sub>I</sub>] cannot, according to the skeptic, serve as the sought after meaning-constituting facts. So it would be very odd if the conclusion of the skeptical argument was that the sought after meaning-constituting facts are facts relating me to my community. Thus, it appears that [SC] cannot be interpreted as [d].

Of course, it doesn't follow (nor am I suggesting that it does) from the claim that facts relating me to my community cannot serve as a direct solution to the skeptical problem, that Kripke also rejects a weaker form of anti-individualism, according to which meaning-ascriptions of the form 'It is true that Jones means such-and-such' or 'It is a fact that Jones means such-and-such' are, in some loose way, *justified* or made *legitimately assertible* by facts about our linguistic community. As will become clear in §1.3.2, this very thought forms the basis of his interpretation of Wittgenstein's private language argument. If an agent's linguistic behavior cannot be evaluated with respect to some

---

speaker's linguistic dispositions and the community's stable linguistic practices. [2008 p. 4] This point seems off to me, for what the skeptical conclusion (or solution) is denying is *not* that these relations can be thought of as stating facts, but that they can serve as the truth conditions for meaning ascriptions.

linguistic community or other, we are not justified in asserting that he means anything at all. Though the brand of anti-individualism captured in [d] isn't up to the task of offering up a direct solution to the skeptical problem, it doesn't follow (as we shall see in §1.3.2) that it isn't up to the task of serving as part of a skeptical solution to the skeptical problem. On the contrary: it is integral to Kripke's skeptical solution to the skeptical problem.

*e. Anti-reductivism*

The most popular contemporary reading of the skeptical conclusion is the following<sup>43</sup>:

- [e]     Meaning-ascriptions cannot be analyzed in terms of any collection of natural, non-semantic facts.

On this reading, [e] is supposed to be what the skeptic means when he denies that there is a 'superlative fact' that corresponds to an agent's meaning so-and-so by w. [e] has been expressed in a variety of ways: there are no meaning FACTS; there are no 'genuine' meaning facts; there are no 'robust' meaning facts, etc. On this reading, the skeptical conclusion is intended to deny the reducibility (in any sense) of meaning attributions to

---

<sup>43</sup> It is defended by, e.g., Horwich [1998], Chapter 10, amongst others.

non-semantic statements.<sup>44</sup> So the moral to pull out of the skeptical argument is that the search for a reductive analysis of meaning-ascriptions is misbegotten. For the most part, the concepts that philosophers single out for attention, like reality, truth, knowledge, beauty, goodness, space, time, and so on, are the most elementary concepts we have. Why, then, one might wonder, should we expect to be able to reduce these concepts to others that are simpler, clearer, and more basic? Donald Davidson raises this question in ‘The Folly of Trying to Define Truth’, and concludes that ‘what makes these concepts so important must also foreclose on the possibility of finding a foundation for them that reaches deeper into bedrock’.<sup>45</sup> Perhaps the skeptic is drawing the same conclusion about meaning-ascriptions: that it is folly to attempt to produce correct and revealing definitions of them in terms of clearer or more fundamental concepts; that it is folly to search for a significant analysis in other terms, for we cannot underpin them with something more basic or easier to grasp. This does not mean that meaning-ascriptions are mysterious, ambiguous, or untrustworthy. Nor does it mean that we cannot say anything revealing about them: we can, by describing how they are *used*; how they *function* in our language.

The following passage might be taken as evidence for this interpretation:

It is important to realize that we are **not looking for necessary and sufficient conditions (truth-conditions)** for following a rule, or an **analysis** of what such rule-following ‘consists in’. Indeed such conditions would constitute a ‘straight’ solution to the sceptical problem, and have been rejected. [p. 87]<sup>46</sup>

---

<sup>44</sup> As Paul Horwich puts it [1998 p. 216]. He also describes it as follows: ‘The thesis is that meaning attributions are not made true by any conglomeration of non-semantic phenomena – so they cannot be incorporated into any grand unified theory of the natural world. [1998 p. 216]

<sup>45</sup> ‘The Folly of Trying to Define Truth’, reprinted in Blackburn and Simons (eds) *Truth*, Oxford, 1999. p. 309

<sup>46</sup> According to Horwich [1998 p. 216-7], additional evidence is to be found by the shape of the skeptical argument: ‘the strategy of the sceptical paradox is to consider a series of non-semantic candidates for the

As in this passage, Kripke often uses the notions of ‘truth condition’, ‘definition’, and ‘analysis’ interchangeably. Thus, he also describes the skeptical conclusion as follows: There are no ‘facts’ or ‘truth-conditions’ [p. 72, 77] or ‘conditions in the world’ [p. 72, 86] corresponding to meaning-ascriptions.

In denying that meaning-ascriptions have truth-conditions, neither the skeptic nor KW (who accepts the skeptical conclusion) wish to deny that meaning-ascriptions have truth-conditions in a ‘minimal’ sense, according to which for any sentence *S* that expresses a proposition, *S* iff it is true that *S* iff it is a fact that *S*. For we already saw that both the skeptic and KW regard the minimal sense of ‘truth’ and ‘fact’, as prefixed or suffixed to meaning-ascriptions, as unproblematic.

Though I believe [e] approximates the content of the skeptical conclusion more accurately than any other candidate we have looked at, I nevertheless do not believe [e] is the correct interpretation of the skeptical conclusion. The quoted passage from WRPL is a bit misleading because a ‘straight solution’ is *not merely* one that offers an analysis of meaning so-and-so by *w*. Any fact that figures in such an analysis must *also* satisfy [J<sub>1</sub>]. Thus, if we take [e] to mean or to leave open the possibility that there is a ‘state’ that satisfies [J<sub>1</sub>], and that it is non-analyzable/primitive, then [e] is clearly not what the skeptic intends by [SC]. Recall that after examining a number of candidate meaning-constituting facts and finding them inadequate (because of their failure to satisfy [J<sub>1</sub>]), Kripke’s skeptic turns to the proposal that meaning so-and-so by *w* is a non-analyzable/primitive state: a state ‘not to be assimilated with sensations or headaches or

---

facts in virtue of which “plus” means what it does, and to argue against each of them, concluding that there are no facts of the right sort to identify with facts about meaning.’

any “qualitative” state, nor to be assimilated to dispositions, but a state of a unique kind of its own’. But he finds this proposal inadequate as well, *for the same reason* that he found the candidate straight solutions inadequate: for its failure to satisfy [J<sub>1</sub>]. What this suggests is that the proper way to interpret [SC] is as follows:

[f] There is no state of an agent, analyzable or primitive, that satisfies [J<sub>1</sub>].

Moreover, [e] cannot be the proper way to interpret the skeptical conclusion because it simply does not follow from the skeptical argument. If there are no facts that satisfy [J<sub>1</sub>], then it does not follow that there are no facts that constitute a reductive analysis for meaning ascriptions.

We can offer an even *more accurate* interpretation of the skeptical conclusion than [f]. Recall, according to my analysis of [J<sub>1</sub>] in §1.1:

[metaphysical constraint] + [epistemological constraint] = [justification constraint [J<sub>1</sub>]]

**Metaphysical constraint** (concerning the normativity of meaning): F contains a norm, a standard of correctness, which determines in advance how w is to be applied in an open-ended range of situations.

**Epistemological constraint**: I have direct access to F (so that it can guide future applications of w).

**Justification constraint [J<sub>1</sub>]**: F justifies my applying word w thus-and-so, given what I mean by it, if and only if (1) I have *direct (non-inferential) access* to F, and (2) I can *deduce a priori* from F that I ought to apply w thus-and-so.

The *key* ingredient of [J<sub>1</sub>] that all facts discussed in the skeptical argument fail to meet is the metaphysical constraint (candidates [A] and [B] do, at least, satisfy the epistemological constraint). So an even more accurate way of interpreting the skeptical conclusion is:

[f\*] There is no state of an agent, analyzable or primitive, that ‘contains’ a norm, a standard of correctness, which determines in advance how w is to be applied in an open-ended range of situations.

What I take the skeptic to be ultimately attacking is the idea that there are states containing instructions on how to apply a word in any conceivable situation. Whether these states are reducible or irreducible, speaker dependent or speaker independent, is not what is ultimately at issue, but their very ability to pull off this feat. There are no states that have this property, according to the skeptical argument. *That* is what it purports to show, and that is how I interpret [SC].

In surveying the literature, I have found that there is at least one commentator – George Wilson -- who agrees that this is the proper way to interpret a fundamental premise of the skeptical argument, if not the skeptical conclusion, as I do here; a view that he also attributes to KW, as I do. On Wilson’s interpretation [1994], the skeptical argument runs as follows [my summary]:

*Wilson’s interpretation of the skeptical argument*

- [N<sub>s</sub>] (Normativity constraint) If X means something by a word w, then there is a set of properties P<sub>1</sub>...P<sub>n</sub> that govern the correct application of w for X. I.e., there is a set of properties P<sub>1</sub>...P<sub>n</sub> such that X has the intention (has accepted the linguistic commitment) that w correctly applies to an object o iff o has P<sub>1</sub>...P<sub>n</sub>.<sup>47</sup>
- [G] (Grounding constraint) If there is a set of properties P<sub>1</sub>...P<sub>n</sub> that govern the correct application of w for X, then there are facts about X that constitute P<sub>1</sub>...P<sub>n</sub> as the conditions that govern X's use of w. I.e., there must be psychological and/or social facts about X which establish that it is P<sub>1</sub>...P<sub>n</sub> that she has successfully singled out and about which she has formed a proper semantic commitment.

---

<sup>47</sup> Scott Soames points out [1998 p. 329ff] that there is a way of interpreting [N<sub>s</sub>] such that it can be seen as a *consequence* of the minimal truth-conditional conception of meaning and understanding (henceforth [MTC]), which he describes, for predicates, as follows: 'A speaker understands a predicate φ that expresses the property of being F iff the speaker knows that φ expresses the property of being F, and hence knows that φ (correctly) applies to an object o iff o is F'. But, Soames argues, [MTC] is virtually undeniable:

It is obvious that sentences and predicates do have truth and application conditions of this sort, and speakers do have the relevant beliefs and intentions. Try this for yourself. All of you who understand the predicate 'red' (and also what it means for a predicate to apply to an object), ask yourselves whether you know that the predicate 'red' correctly applies to an object iff it is red. The answer, of course, is that you do. The skeptical solution cannot plausibly be seen as denying this triviality.

Consequently, Soames argues, one must come up with a stronger, more substantial, and more plausibly deniable interpretation of 'classical realism about truth-conditions'. [1998 p. 330] On behalf of Wilson, Soames offers the following strengthened account of [N<sub>s</sub>]:

- [N<sub>s</sub>+] If x means something by word w, then x does so **because** there is a set of properties P<sub>1</sub>...P<sub>n</sub> of which x has **prior grasp**, and x has the intention that w correctly applies to an object o iff o has the properties P<sub>1</sub>...P<sub>n</sub>.

What the skeptic is rejecting, on this reading of governance is that a speaker's understanding of a word, including his ability to apply it to newly considered objects, is explained by his associating w with independently apprehended properties P<sub>1</sub>...P<sub>n</sub>, and judging those objects to have P<sub>1</sub>...P<sub>n</sub>.

[BSC] (Basic skeptical conclusion) There are no facts about X that constitute any set of properties as conditions that govern X's use of w. I.e., there are no facts in the world to be correctly recorded by ascriptions of meaning.

∴ [RSC] (Radical skeptical conclusion) No one ever means anything by any term.

Wilson maintains that unlike Kripke's skeptic, Kripke's Wittgenstein sides with the common opinion that people often mean something by the terms and sentences they use. He thereby rejects [RSC]. But he purports to have shown that those terms and sentences cannot have 'classical realist' truth and application conditions, so that the meaningfulness of language for its users cannot consist in that. Thus, he accepts [BSC]. And hence, he is committed to rejecting [N<sub>S</sub>]. [1994 pp. 374]

In sum, on Wilson's reading, the difference between Kripke's skeptic and Kripke's Wittgenstein is that:

<u>Kripke's skeptic</u>	<u>Kripke's Wittgenstein</u>
N <sub>S</sub>	G
G	BSC
BSC	~RSC
∴ RSC	∴ ~N <sub>S</sub>

On the interpretation that I have offered of the skeptical argument, I do not see the radical skeptical conclusion [RSC] as dropping out of it, as Wilson does. If we can drive a wedge between Kripke's skeptic and Kripke's Wittgenstein, it is not by attributing the

radical skeptical conclusion to Kripke's skeptic and dissociating it from KW. The difference can be framed, rather, through [J]: KS insists that this constraint be met by any fact that corresponds to a subject's meaning what he does by his words, whereas KW rejects it.

But what, according to Wilson, is KW rejecting by rejecting [BSC]? In a recent paper, Wilson argues that what KW is rejecting is the notion of 'prior semantic determination' or 'factual determinacy' about 'pre-established standards of correctness' [2008 p. 7]. This is what the 'classical realist picture of meaning' amounts to: the idea that 'what a speaker means now by a term determines how the term, in its present meaning, is to be applied correctly in an indefinite range of yet to be examined cases' [2008, p. 1]; 'that the meaning of a term semantically determines in advance whether or not the term, so meant, applies to various actual or possible candidate items' [2008 p. 1-2]. Wilson takes it, correctly in my opinion, to be a 'great achievement of Kripke's book to stress the centrality of the skeptical doubts about semantic determination in Wittgenstein's rule following reflections and to elaborate an original and powerful challenge to this fundamental idea' [2008 p. 2]. Unfortunately, he notes, this achievement has been blurred by 'Kripke's repeated invocation of a striking "nonfactualism" about meaning': the claim that there are no facts, individualistic or social, about a speaker that constitute his meaning one thing rather than another (and, rather than nothing at all) by his use of any given term or sentence. I agree with Wilson on both of these points, and, as I try to show in chapter 4, skeptical doubts about prior semantic determination do indeed play a central role in Wittgenstein's discussion of meaning and understanding.

It is generally agreed that if a subject means something by a word *w*, there must be standards that determine whether he applies *w* correctly, and his use of *w* must conform to those standards if it is correct. This is what it means for meaning to be *normative*. But it is easy to slip into thinking that these standards are fully determined in advance, that they settle every case, and that they are unchanging, as opposed to being constructed over time, and hence also, alterable over time. In §2.4, I will show that we have good reason to think that this is the true nature of meanings: they are not stagnant, like dead matter, but alter and adapt to their environment with the passing of time, like living organisms. Those that fail to do so simply cease to exist, or exist only in the space of conceptual possibilities. Nietzsche's discussion of the origin of 'good' and 'bad' illustrates this point well. Many centuries ago, "good" represented the dominant temper of a higher, ruling class in relation to a lower dependent one.' It is only after the aristocratic values began to decline that 'the egotism-altruism dichotomy' took 'possession of the human conscience'<sup>48</sup>, and 'good' acquired a new meaning. Such an origin suggests, to Nietzsche, that there is 'no a priori necessity for associating the word good with altruistic deeds'<sup>49</sup>, as we do in contemporary culture. The old meaning was no longer functional in a changing society, and was gradually replaced with a new one. The notion of 'democracy', which altered from meaning direct rule to indirect rule by the people, serves to illustrate this point just as well. The platonist picture, which conceives of meanings as eternal, unchanging, and fixed in advance for all time, is necessarily flawed. It was, of course, a central preoccupation of Wittgenstein's to dismantle it, in all its beguiling

---

<sup>48</sup> Francis Golffing (tr.) *The Birth of Tragedy and the Genealogy of Morals*, Doubleday Press, 1956, pp. 160-161.

<sup>49</sup> *Ibid*, p. 160.

forms. And it is an important though unrecognized merit of Kripke's book to recognize this as the moral of his rule-following discussion.

### 1.3.2 Kripke's Interpretation of the Private-Language Argument

Kripke seeks to make his interpretation of Wittgenstein's position – as advancing a 'skeptical solution' to the skeptical problem -- plausible by linking it to Wittgenstein's abandonment of the 'classical realist' theory of meaning found in the *Tractatus* in favor of the conception of meaning as *use* advocated in his later writings. [p. 71ff]

Whether or not Wittgenstein is right in thinking that the entire *Tractatus* view is a consequence of natural and apparently inevitable presuppositions, he is surely right about this fundamental part of it. The picture of **correspondence to facts** must be cleared away before we can deal with the skeptical problem. [p. 78-9]

Wittgenstein's skeptical solution concedes to the skeptic that no '**truth-conditions**' or '**corresponding facts in the world**' exist that make a statement like 'Jones, like many of us, means addition by "+" true. [p. 86]

According to Kripke [p. 73], Wittgenstein replaces the question of the *Tractatus*:

- i. What must be the case for the sentence 'Jones means addition by "+" to be **true**?

by two others in the *Philosophical Investigations*:

- ii. Under what conditions can we **justifiably assert** (or deny) that Jones means *addition* by ‘+’?
- iii. What is the **role and utility** in our lives of our practice of asserting (or denying) such statements?

Kripke interprets Wittgenstein as offering the view that I summed up as [J<sub>E</sub>] in answer to (ii): ‘Jones means *addition* by “+”’ is warranted/justified iff Jones’ use of ‘+’ matches the use of ‘+’ in his linguistic community, and his linguistic community means *addition* by ‘+’. And in answer to (iii), Kripke interprets Wittgenstein as offering the view that, in making such statements, we are not *representing* the world as being a certain way, even though we attach notions of truth and falsity to them. Rather, we are *expressing our confidence* in a subject: we are labeling him as someone whose statements involving ‘+’ ought to be relied upon. If I say of a grocer, for example, that he means *addition* by ‘+’, I am saying that he can be trusted not to cheat you when you buy goods at his shop, and if I assert that he does not mean *addition* by ‘+’ (very often can be said of NY grocers), I am warning you against buying goods from his shop. It is in ways like this that meaning-ascriptions have a *role* and *utility* in our lives.

Though KW accepts the conclusion of the skeptical argument ([f] above), he rejects the manner in which it is reached. For the skeptic, the failure of any candidate meaning-constituting fact to satisfy [J<sub>I</sub>] entails the skeptical conclusion. KW, however, rejects [J<sub>I</sub>].

[J<sub>I</sub>] F justifies my applying word w thus-and-so, given what I mean by it, if and only if (1) I have *direct (non-inferential) access* to F, and (2) I can *deduce a priori* from F that I ought to apply w thus-and-so.

He concedes that ‘our ordinary practice or belief is justified because – contrary appearances notwithstanding – it need not require the justification that the skeptic has shown to be untenable’. [p. 67]

[J<sub>I</sub>] is an *internalist* account of justification, which makes reference only to an individual. KW insists that the failure of any fact to satisfy [J<sub>I</sub>] does not undermine our ordinary practice of meaning-attributions, because that practice can be justified in another, more plausible sense.

Any individual who claims to have mastered the concept of addition will be judged **by the community** to have done so if his particular responses **agree with those of the community** in enough cases, especially the elementary ones (and if his ‘wrong’ answers are not *bizarrely* wrong, as in ‘5’ for ‘68 + 57’, but seem to agree with ours in *procedure*, even when he makes a ‘computational mistake’). An individual who passes such tests is admitted into the community as an adder; an individual who passes such tests in enough other cases is admitted as a normal speaker of the language and member of the community. Those who deviate are corrected and told (usually as children) that they have not grasped the concept of addition. One who is an incorrigible deviant in enough respects simply cannot participate in the life of the community and in communication. [p. 91-92]

I summed up the more plausible sense in which our linguistic practice is justified, according to KW’s skeptical solution, as:

[J<sub>E</sub>] ‘S means M by w’ is warranted/justified iff S’s applications of w accord with those of his linguistic community, and M is the meaning w has in S’s linguistic community.

In its application to the private language argument, [J<sub>E</sub>] takes a stronger form:

[J<sub>E+</sub>] ‘S means M by w’ is warranted/justified iff (a) S’s applications of w accord with those of his linguistic community,<sup>50</sup> and M is the meaning w has in S’s linguistic community, and (b) it is possible to check that S’s applications of w accord with those of his linguistic community.

Unlike [J<sub>I</sub>], which grounds justification-conditions for a subject’s applying a word in facts *internal* to the subject, [J<sub>E+</sub>] grounds justification-conditions for a subject’s applying a word in (social) facts *external* to the subject – specifically, in communal practices involving that word. Justification conditions, according to KW, cannot be internal, as the skeptical argument has presumably established, but must involve reference to others: the distinction between correct (in the sense of ‘warranted’) and incorrect rule-following must be framed in terms of agreement or absence of agreement with communal practices.

The impossibility of a private-language, Kripke maintains, emerges as a ‘corollary’ to the skeptical solution. [p. 68] Kripke discusses the relation between the skeptical solution and the private language argument in the following passages.

---

<sup>50</sup> We could generalize this to ‘some linguistic community’, rather than S’s. S’s behavior would then have to accord with the behavior of *that* linguistic community.

If one person is **considered in isolation**, the notion of a rule as guiding the person who adopts it can have no substantial content. There are, as we have seen, no truth-conditions or facts in virtue of which it can be the case that he accords with his past intentions or not. **As long as we regard him as following a rule “privately”, so that we pay attention to his justification-conditions alone, all we can say is that he is licensed to follow the rule as it strikes him.** This is why Wittgenstein says, ‘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.’ (§202) [p. 89]

Following §243, a private-language is usually defined as a language that would be logically impossible for anyone else to understand. The private-language argument is taken to argue against the possibility of a private-language in this sense. This conception is not in error, but it seems to me that the emphasis is somewhat misplaced. **What is really denied** is what might be called the ‘**private model**’ of rule-following, that **the notion of a person following a given rule is to be analyzed simply in terms of facts about the rule follower and the rule follower alone, without reference to his membership in a wider community.** The impossibility of a private-language in the sense just defined does indeed follow from the incorrectness of the private model for language and rules, since the rule-following in a ‘private-language’ could only be analyzed by a private model, but the incorrectness of the private model is more basic, since it applies to all rules. I take all this to be the point of §202. [p. 109-110]

We can sum up Kripke’s interpretation of the private language argument [KPLA] as follows:

*Kripke’s interpretation of PLA*

- i. Language is a rule-governed activity, and it can only exist where it makes sense to say someone is following a rule. (Here Wittgenstein’s comparison of language to a game is pertinent.)
- ii. (Corollary to the skeptical solution) It makes sense to say someone is following a rule (applying a term ‘correctly’) *only if* there is a way for people

to tell whether the rule has been followed or broken. For this to be accomplished:

(a) The standards for what counts as a correct application must be *external* to the subject. More specifically, they must be grounded in communal/social practices.<sup>51</sup>

(b) The subject's applications of *w* must (in principle) be *checkable* by others, who compare S's applications of *w* to the communal standards mentioned in

(a). It is only they, and not S alone, who are in a position to tell whether S is following a rule. S may think he is following a rule when in fact he is not: there is a difference between thinking one is following a rule and following the rule.

iii. A 'private language' is by definition one that is 'logically impossible' for anyone other than the subject S to understand. Hence, in a purported private

---

<sup>51</sup> Colin McGinn [1984] argues that the reference to the community is redundant in the justification condition 'for the correct condition is simply that I observe that you give the sum of pairs of numbers sufficiently often. This conclusion is of course entirely individualistic in that it refers only to the person to whom the rule is ascribed and to his behavior' [1984 p. 185] 'It is not essential to make reference to the community in giving the criteria for (e.g.) meaning addition by "+" [1984 p. 187] The point is remade in Goldfarb [1985 p. 482 ff]. Why, Goldfarb argues, can't we simply frame the assertibility/justification conditions as follows:

(\*) It is licensed to assert that a person means addition by '+' when that person has responded with the sum in every case so far attempted.

Goldfarb continues: 'Kripke's objection was that, in the solitary case, there would be no difference between an individual's actually fulfilling a condition and its appearing to him that he fulfills it. Clearly however, with respect to (\*) there is a difference. In short, descriptively adequate assertibility conditions can be framed using (not mentioning) what is ascribed as being meant, that is, notions like addition, table, green, chair, and so on. Such conditions invoke neither community nor agreement. Yet they do yield a distinction between an individual's fulfilling a condition and the individual's taking himself to fulfill it, and they do provide a basis for distinguishing between responses that are in accord with a previously ascribed meaning and responses an individual is simply inclined to make. Thus Kripke's central claim, that his solution shows how public language is impossible but solitary is not, collapses.' [1985 p. 483]

In chapter 2.1, I will evaluate these claims (as they are made within the context of Kripke's dispositional theory), and show that they commit us to an implausible platonist metaphysics of meaning.

language, no one is ever in a position to tell whether S is applying his words correctly, since no one is ever, even in principle, able to check his applications.

- iv. So, it never makes sense to say someone is following a rule in a purported ‘private language’
- v. So there cannot be (conceptually speaking)<sup>52</sup> a private language.

The justification conditions for when a word *w* has been applied ‘correctly’ (for *any* word *w*) must be speaker independent / external to the agent in the two crucial senses captured in (a) and (b).

It is worth noting that (b)  $\leftrightarrow$  (a). As Colin McGinn points out:

‘What we need for a genuine community conception is the idea that following a rule is a notion, like that of being in fashion or marching in step, that inherently involves reference to individuals other than the individual to whom the ascription is made: and the requirements of third person-knowability and of sharability do not entail anything of that kind. Neither does the idea that rule-following has to be in principle correctable by others, if indeed this goes beyond knowability and sharability.’ [1984 p. 194]

It is also worth noting that (a)  $\leftrightarrow$  (b). The standards for what counts as a correct application of *w* may be grounded in communal practices, even though a subject’s use of *w* is not in principle checkable by others. Suppose e.g. S uses the term ‘red’ to represent a private sensation that looks like the color red. So long as his applications conform to

---

<sup>52</sup> There ‘cannot’ be a private language in the same sense that, for example, Thrasymachus in Book I of the *Republic* says rulers ‘cannot’ make mistakes. Insofar as S is a ruler (i.e., *qua* ruler), he cannot make mistakes. This is a conceptual point about rulers. Thrasymachus is not saying that people we call rulers do not in fact make mistakes.

communal applications of 'red', one is justified in asserting that 'S means red by "red"', according to (a), even though S's applications of 'red' are not in principle checkable by others. Thus, the conclusion of [KPLA] cannot follow from (a) alone, i.e., from [J<sub>E</sub>]. Only if we add (b) to [J<sub>E</sub>], as we have done in [J<sub>E</sub>+], do we exclude the possibility of a 'private language'.

[KPLA] is stronger than typical *verificationist* interpretations of the private language argument which require (b) alone, for, according to Kripke, even if (b) is satisfied, a 'private language' may still be impossible. In addition to (b), there is another sense in which, according to Kripke, the standards for what counts as a 'correct' application must be speaker independent: they must be grounded in communal practices, for *it is only in communal practices that standards can be created*. On Kripke's reading of the private language argument, 'means so-and-so' functions like the *relational* predicate 'tall' T(x,y) where  $y \neq x$ . When we say x is tall, we are comparing him to some person(s): y. x may be tall relative to y' and short relative to y''. Likewise, 'means so and so' is a relational predicate M(x,y) where  $y \neq x$ . 'Means so and so' or 'follows such and such a rule' are, for Kripke, notions like that of being in fashion or marching in step that essentially involve reference to others, to use McGinn's examples [1984 p. 194]. When we say x means so-and-so by w, we are comparing his use of w to that of some linguistic community y. It is regularities in the collective behavior of y that serve as the comparative base, and determine the norms of correct usage. Regularities in the behavior of a *single* individual cannot generate norms. They cannot constitute a 'practice', as Kripke interprets Wittgenstein's use of this expression in PI §202:

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.

Meanings, on this picture, emerge in social practices, in more or less uniform collective facts. These are the surroundings (y) where standards (meanings) are created, and an individual x can only be said to use language meaningfully when his behavior is held accountable to these standards.

Kripke notes that it does not follow from [KPLA] that Robinson Crusoe, isolated on an island, cannot be said to follow rules. [p. 110]

If one person is **considered in isolation**, the notion of a rule as guiding the behavior of the person who adopts it can have *no* substantive content. ... The situation is very different if we widen our gaze from consideration of the rule-follower alone and allow ourselves to **consider him as interacting with a wider community**. Others will then have the justification conditions for attributing correct or incorrect rule-following to the subject, and these will not simply be that the subject’s own authority is unconditionally to be accepted. [p 89]

What follows from [KPLA] is that:

if we think of Crusoe as following rules, we are taking him into our community and applying our criteria of rule-following to him. The falsity of the private model need not mean that a *physically isolated* individual cannot be said to follow rules; rather that an individual, **considered in isolation** (whether or not he is physically isolated), cannot be said to do so. [p. 110]

The point is that [KPLA] does not require Crusoe’s *actual membership* in a linguistic community for others to be warranted in ascribing rule-following to him<sup>53</sup>. [KPLA]

---

<sup>53</sup> Note also that [KPLA] does not require Crusoe’s actual membership in a linguistic community for him to follow rules. It only addresses what we are or are not *warranted* in asserting about the solitary ‘rule-follower’. Thus, McGinn’s objection to [KPLA], that ‘it would, I think, be totally implausible to maintain

requires only that they be able (in principle) to compare his behavior to that of some linguistic community. We would, e.g., be entitled to say that Crusoe follows the ‘plus’ rule if he applies some symbol the way *we* apply ‘+’. Agreement with the standards of a community that follows the ‘+’-rule is essential if the ‘+’-rule is to be properly ascribed to Crusoe.

We have seen that it is part of Wittgenstein’s *general* view of the workings of *all* our expressions attributing concepts that others can confirm whether a subject’s responses agree with their own. [p. 100]

Nor does [KPLA] require that there be others who follow the *same* rule as Crusoe. For Crusoe’s rule might be defined in terms of simpler rules that others follow. Suppose, e.g., that after long observations of Crusoe’s behavior, we form the hypothesis: ‘Crusoe uses “blurb” to refer to what we take to be red or orange objects’. Then we are entitled to say ‘Crusoe means *red or orange* by “blurb” so long as he applies ‘blurb’ to what *we* take to be red or orange objects. Thus, [KPLA] does not rule out the possibility of ascribing rule following to, e.g., a creative mathematician who discovers a function which he names, or a zoologist who comes across a hitherto unknown species which he names.<sup>54</sup> For presumably these new expressions are or can be defined in terms of others that already possess a meaning in the agent’s linguistic community, so that whether he is applying the terms in question correctly can be reduced to the question of whether he is

---

that Crusoe somehow *loses* his rule-following capacities when he wakes up on his island in social isolation, that his words no longer have meaning (what if you went alone to a desert island for a holiday?)’ [1984 p. 195] misses the point of [KPLA].

<sup>54</sup> As Colin McGinn seems to think. [1984 p. 195]

applying the more basic expressions in accord with how his linguistic community applies them.

Still, [KPLA] leaves important questions unanswered. Suppose a subject speaks a language shared by *no* community, including ours, and has never had contact with any rule-followers. The case is described by McGinn.

[Suppose] Romulus, upon reaching the age of reason, hits upon the idea of distributing sign-posts around his island as an *aide-memoire*. He wants to avoid the marshes, so he writes an arrow in the sand and undertakes to walk in the direction of its head when he comes across it in the future; he follows his rule correctly in the future if he conforms his actions to his original intentions in respect of the arrow; and he may discover on occasion that he has followed his rule incorrectly when, misremembering his original intention, he finds himself wallowing in the marsh (he mistakenly thought that his original intention was to follow the *tail* of the arrow). Nor do I see any conceptual obstacle to his introducing properly linguistic signs for his own use, e.g. to keep records of the weather: all he needs is a good reason to introduce the signs and the intelligence to operate with them. That he has never had contact with other rule-followers does not seem to me to put a logical wall between him and the activity of rule-following. Certainly it seems implausible to insist that rules must necessarily be *taught* if they are to be grasped: for who taught the teachers? [1984 p. 195-6]

Chomsky offers the following account of how Kripke would respond to this case:

‘Kripke argues that we can still attribute rule-following to RC in accordance with the Wittgensteinian solution. Namely, we regard him as a person who acquires rules under certain experiences, although not our rules, because we had different experiences. Then we can take him into the broader community of persons, who share our ‘form of life’ in a broader sense. “Our community can assert of any individual that he follows a rule if he passes the tests for rule-following applied to any member of the community,” that is, if he acts in the manner of rule follower although he doesn’t give our responses.’ [1986 p. 230-231]

In order for us to attribute personhood to a subject, Chomsky maintains, he must exhibit ‘characteristic features of will and choice, the creative aspect of language use, and other

indications of intelligence'. [1986 p. 235. See also p. 230] However, I see no textual evidence for the claim that Kripke uses the attribution of personhood as a criterion for the attribution of rule-following. Rather, I think Kripke would deal with this case in the same way as the Crusoe case above. So long as Romulus' behavior exhibits observable (in principle) regularities that mimic the regularities of some linguistic community, we are justified in ascribing rule-following to him. This does not imply that:

- (i) Romulus must be a member of a linguistic community.
- (ii) Romulus was once a member of a rule-following community.

But [KPLA] does appear to imply that:

- (iii) Other language users must *actually* exist that serve as the comparative base (y).

According to McGinn, this presents a serious problem for [KPLA], for surely 'Adam [did not have to] wait for Eve before he could follow rules. ... How could the creation of other individuals wreak such a momentous change in the first created individual?' [1984 p 191 n] '... could not the rest of the human race be wiped out while you sleep and yet the next day you awake with your rule-following capacities intact?' [1984 p. 198]

It seems to me that this objection misses the point of [KPLA], for the conclusion of [KPLA] does not concern what is / is not true of Adam, but what we are / are not entitled to *say* about him. We are not entitled to say Adam, or the last human on the planet, are

following rules, because there is no one to compare their behavior to.<sup>55</sup> What follows from this, according to [KPLA] is that we cannot *call* their linguistic behavior a language; not that they cannot engage in the kind of behavior they do.

Kripke points out that a community's means of checking whether a member applies terms as they do will vary over different discourses. For example, in the case of a term like 'table', which refers to observable physical objects, we judge a person to have mastered the term if he applies it to objects in our visual field which we agree are tables. In the case avowals of pain (e.g., 'I am in pain'), we obviously can't observe a person's sensation of pain; instead, we check to see if he exhibits what we take to be pain-behavior when he uses the avowal. Likewise, in the case of mathematical statements, we don't compare a person's use of them to some supra-sensible reality, but to our own use of them. In general, according to [KPLA], we are justified in attributing meaning to an individual to the extent that we can compare his behavior to our own.

---

<sup>55</sup> It might be thought that KPLA implies a weaker thesis:

- (v) A *possible* or *hypothetical* community of language users must exist that serves as the comparative base (y).

To illustrate, at a certain stage of language growth, as Chomsky observes, children characteristically overgeneralize: they say *sleeped* instead of *slept*, *brang* (on analogy with *sang*) instead of *brought*. Now in fact there is no community that generalizes in this way. But we can imagine our community doing so. We would then be employing the child's 'rule', and we can say his applications are correct if he generalizes the way *we* would generalize in such a hypothetical community. But I have not been able to find textual evidence that this is what Kripke intended.

## 2

## Evaluating the Skeptic's Arguments Against the Dispositional Theory of Meaning

---

### CONTENT

2.0	Introduction
2.1	Overcoming the Finitude Objection
2.1.1	The Finitude Argument
2.1.2	Critique of the Finitude Argument
2.2	Overcoming the Error Objection
2.2.1	The Error Argument
2.2.2	Critique of the Error Argument
2.3	Overcoming the Circularity Objection
2.3.1	The Circularity Argument
2.3.2	Critique of the Circularity Argument
2.4	Overcoming the Justification Objection
2.4.1	The Justification Argument
2.4.2	Critique of the Justification Argument
2.5	Conclusion

---

### 2.0 INTRODUCTION

In Chapter 2 of WRPL, Kripke raises the following (my summary) objections against the dispositional theory of meaning (applied to the term '+'):

- (1) **Finitude objection:** Since my actual dispositions are *finite*, it's not true that if queried about any two numbers I will come up with their sum, as the dispositional theory maintains, for some numbers are too large for my mind to

grasp. But surely I can mean *addition* by ‘+’ despite the fact that I don’t have dispositions to handle huge numbers.

- (2) **Error objection:** Any meaning-constituting fact must be able to account for the fact that some people are disposed to make *mistakes*. But the dispositional theory cannot account for this: according to it, how one *should* apply ‘+’ is how one *would* apply ‘+’.
- (3) **Circularity objection:** The dispositional theory cannot succeed in giving a non-circular description of the facts that constitute my meaning *addition* by ‘+’.
- (4) **Justification objection:** Any meaning-constituting fact must ‘show how I am justified’ [p. 11] in giving one answer rather than another to any particular addition query. I.e., any meaning-constituting fact must be such that, however in fact I am disposed to apply ‘+’, it should ‘tell me’ [p. 24] how I *should* apply it. But dispositions do not have this property.

These objections have a common form (as will become apparent when we look at them in detail below): they impose a specific constraint on the dispositional theory, and then reject it as a straight solution to the skeptical problem

[SKEP] Is there any fact about me that constitutes my meaning *addition* by ‘addition’ (or, more generally, my meaning *w* by ‘w’, for any expression *w* of our language)?

for failing to meet that constraint.

In this chapter, I examine and critique each objection. The chapter is divided into four sections, corresponding to the four objections. The plan for each section is as follows.

- i. I begin by laying out the argument corresponding to the objection in question, so as to identify the constraint that the objection imposes on the dispositional theory.
- ii. I then argue that the argument's success depends on an unwarranted assumption Kripke makes about how the constraint is to be interpreted, and that no meaning-constituting fact should be required to meet the constraint thus interpreted.
- iii. Finally, I offer a more reasonable interpretation of the constraint in question, and try to show that a more refined version of the dispositional theory than that entertained by Kripke is invulnerable to the objection when the constraint is thus interpreted.

I develop this more refined version of the dispositional theory – what I will call ‘**constructive dispositionalism**’ -- in piecemeal fashion throughout this chapter, as we deal with each objection in turn.

## 2.1 OVERCOMING THE FINITUDE OBJECTION

### 2.1.1 The Finitude Argument

Kripke's finitude argument against the dispositional theory of meaning emerges in the following passage:

The dispositional theory attempts to avoid the problem of the 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. It is not true, for example, that queried about the sum of any two numbers, no matter how large, I will reply with their actual sum, for some pairs of numbers are simply too large for my mind -- or my brain -- to grasp. When given such sums, I may shrug my shoulders for lack of comprehension; I may even, if the numbers involved are large enough, die of old age before the questioner completes his question. Let 'quaddition' be redefined so as to be a function which agrees with addition for all pairs of numbers small enough for me to have any disposition to add them, and let it diverge from addition thereafter (say, it is 5). Then, just as the skeptic previously proposed the hypothesis that I meant quaddition in the old sense, now he proposes the hypothesis that I meant quaddition in the new sense. A dispositional account will be impotent to refute him. As before, there are infinitely many candidates the skeptic can propose for the role of quaddition.

I have heard it suggested that the trouble arises solely from too crude a notion of disposition: *ceteris paribus*, I surely will respond with the sum of any two numbers when queried. ... Perhaps, but how should we flesh out the *ceteris paribus* clause? Perhaps as something like: if my brain had been stuffed with sufficient extra matter to grasp large enough numbers, and if it were given enough capacity to perform such a large addition, and if my life (in a healthy state) were prolonged enough, then given an addition problem involving two large numbers,  $m$  and  $n$ , I would respond with their sum, and not with the result according to some quus-like rule. But how can we have any confidence of this? How in the world can I tell what would happen if my brain were stuffed with extra brain matter, or if my life were prolonged by some magic elixir? Surely such experiments should be left to science fiction writers and futurologists. We have no idea what the results of such experiments would be. They might lead me to go insane, even to behave according to a quus-like rule. The outcome really is obviously indeterminate, failing further specification of these magic mind-expanding processes; and even with such specifications, it is highly speculative. But of course what the *ceteris paribus* clause really means is something like this: If I somehow were to be given the means to carry out my intentions with respect

to numbers that presently are too long for me to add (or to grasp), and if I were to carry out these intentions, then if queried about 'm+n' for some big m and n, I would respond with their sum (and not with their quum). Such a counterfactual is true enough, but it is of no help against the skeptic. It presupposes a prior notion of my having an intention to mean one function rather than another by '+'. It is in virtue of a fact of this kind about me that the conditional is true. But of course the skeptic is challenging the existence of just such a fact; his challenge must be met by specifying its nature. [p. 27-28]

Recall, the problem for which the dispositional theory is entertained as a solution is:

[SKEP] Is there any fact about me that constitutes my meaning *addition* by 'addition' (or, more generally, my meaning *w* by 'w', for any expression *w* of our language)?

Kripke's dispositional theorist offers the following solution to [SKEP]: [pp. 22-3]

[KD<sup>plus</sup>] S's meaning *addition* by '+' consists in the fact that for any p, q, S is disposed to answer r to the query 'p+q=?' iff r is the *sum* of p and q.

One problem Kripke raises for this theory in the quoted passage is that it is not true that if queried about the sum of *any* two numbers, no matter how large, I will answer with their sum: some pairs of numbers are simply too large for my mind to grasp. In this sense, my additive dispositions are *finite* – they can only extend so far. But surely I can mean *addition* by '+' despite the fact that I don't have dispositions to handle huge sums.

This objection shows that the *right to left direction* of the analysis in [KD<sup>plus</sup>] fails, i.e., that:

$\forall p, q, S$  is disposed to answer  $r$  to the query ‘ $p+q=?$ ’  $\leftrightarrow$   $r$  is the *sum* of  $p$  and  $q$ .

and forms the key premise in the finitude argument. As I see it, the argument (which can be extracted from the passage) takes the form of a *reductio*, and may be summed up as follows:

*Finitude argument against the dispositional theory (applied to ‘+’)*

- i Suppose I mean *addition* by ‘+’.
- ii According to the dispositional theory [KD<sup>plus</sup>], I mean *addition* by ‘+’ if and only if I am disposed to give answers that covary with the set of ordered triples that lie in the extension of ‘+’.
- iii Since ‘+’ applies to objects that are *epistemically inaccessible* to me, there will be triples  $\langle p, q, r \rangle$  in the extension of ‘+’ over which I will not be disposed to answer  $r$  to the query ‘ $p+q=?$ ’.
- iv Therefore, according to the dispositional theory [KD<sup>plus</sup>], I do not mean *addition* by ‘+’, or understand it.
- v But this contradicts the supposition (i) that I do mean *addition* by ‘+’. So the dispositional theory [KD<sup>plus</sup>] must be wrong.

It will not help the dispositional theorist to reformulate [KD<sup>plus</sup>], Kripke adds, by appealing to how I would respond under *ideal conditions* that overcome the finiteness:

[KD<sub>r</sub><sup>plus</sup>] To mean *addition* by ‘+’ is to be disposed, *under ideal conditions*, to answer *r* to the query ‘*p+q=?*’ iff *r* is the *sum* of *p* and *q*.

For how are we to flesh out the ideal conditions? The proposal that we characterize them as conditions where ‘I am given the means to carry out my intentions’ will succeed, Kripke says, only if we ‘presuppose a prior notion of my having an intention to mean one function rather than another by “+”’. [p. 28] We cannot confidently say that, if I am given the means to carry out my intentions, I will give answers that accord with the addition function, for how am I to be given these means? If my mind is stuffed with extra brain matter, or my life prolonged by some magic elixir, we would have no idea how I would be disposed to apply ‘+’. Stuffing my brain with extra matter or prolonging my life with magic elixirs might lead me to go insane!<sup>56</sup> And even if I am not led to go insane, there

---

<sup>56</sup> C. B. Martin and J. Heil, in a recent defense of the dispositional theory [1998 p. 300ff], argue that we don’t need to invoke such mind-expanding thought experiments because we *do* in fact possess the ability, and hence disposition (since, according to them, ‘an ability amounts to the possession of a particular sort of dispositional condition’ [p. 301]) to add indefinitely large numbers, but due to limitations on our speed of computation, memory, and assorted physical capacities, we are not able to manifest it. In another possible world, they argue, where there are no such limitations, we are able to manifest it. So it is not that we don’t possess the ability to add indefinitely large numbers, they conclude, but that the manifestation of the ability is blocked or inhibited by various physical limitations.

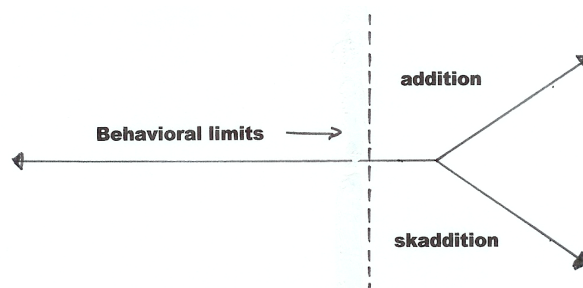
We can sum up their reasoning as follows:

- i. In another possible world where there are no limitations on our speed of computation, memory, and assorted physical capacities, we are able to add indefinitely large numbers.
- ii. Therefore, we *do* in fact possess the ability to add indefinitely large numbers, but the manifestation of this ability is blocked or inhibited by various physical limitations.

This strategy for overcoming the finitude objection, however, is hardly plausible. If this argument was valid, then the following arguments would also have to be valid:

- i. If conditions were ‘ideal’ for a tortoise in another possible world – if, for example, it were much faster than a hare – then it could overtake a hare.
- ii. Therefore, a tortoise has the ability to overtake a hare, but the manifestation of this ability is blocked or inhibited by various physical limitations. [Boghossian’s example from 1989 p. 527]

is no guarantee that I would be disposed to give answers that accord with the addition function as opposed to a function – let us call it skaddition -- that coincides with the addition function on some finite (surveyable) initial segment of the addition function, but that diverges from it thereafter



unless, Kripke says, we presuppose that I have formed an intention *to give answers that accord with the plus function*. But, he maintains, building this clause into the ideal conditions would render  $[KD_r^{\text{plus}}]$  circular.

- 
- i. If conditions were ‘ideal’ for a human in another possible world – if, for example, he had the powers of superman – then he could fly.
  - ii. Therefore, humans have the ability to fly, but the manifestation of this ability is blocked or inhibited by various physical limitations.

Of course, these arguments are invalid. One can hardly credit a tortoise with the ability to overtake a hare, or a human with the ability to fly, by supposing they could in another possible world. Obviously, *only certain idealizations are permissible*. Likewise, I submit, one can hardly credit an agent with the ability to add indefinitely large numbers – numbers that lie beyond our range of comprehension – by supposing that, if in another possible world the agent had no limitations on his speed of computation and assorted physical capacities, he would supply correct answers. Here too, we have transgressed the limits on permissible idealizations. A ‘human’ that has the ability in another possible world to handle indefinitely large numbers – numbers that are too large for even the fastest computer in the actual world to handle -- is *not* human (just as a ‘human’ that has the powers of superman in another possible world is *not* human, or a ‘tortoise’ that overtakes a hare in another possible world is *not* a tortoise). The thought experiment suggests something that is impossible. There is a *limit* on the number of iterations in computations that must be observed in the idealization proposed by Martin and Heil, otherwise it loses its plausibility. The number of iterations must lie *within the bound* of what is humanly possible.

### 2.1.2 Critique of the Finitude Argument

Where, if anywhere, has the finitude argument gone wrong? Premise (iii) assumes that if I mean *addition* by '+', then, I can't be disposed to apply '+' in any old way -- I must apply it correctly. This much seems correct. As Michael Dummett points out, using language is not like doing one's hair and taking a bath. One may do either of these two things as one likes, and still be doing it. But if there cease to be right and wrong uses of a word, the word loses its meaning. [1991, p. 85] But, as Kripke presents the dispositional theory ([KD<sup>plus</sup>]), I am disposed to apply '+' **correctly**, given that I mean *addition* by '+', if and only I am disposed to give answers that covary with the set of ordered triples that lie in the **extension** of '+'. This generates a problem because the extension of '+' involves terms that are ungraspably huge, so that (as premise iii states) the *right-to-left direction* of the biconditional in [KD<sup>plus</sup>] breaks down, i.e.:

$$\forall p, q, S \text{ is disposed to answer } r \text{ to the query 'p+q=?' } \leftrightarrow p+q=r$$

Some terms apply to objects that are **epistemically inaccessible** to us, and hence objects about which we do not have any normal dispositions. 'Plus' is one such term, for a certain range of values. Thus for some trio of integers p, q, and r, such that p+q=r, it is beyond my capacity to consider the question whether p+q=r, and so the statement 'If  $\langle p, q, r \rangle$  is in the extension of "+", I would be disposed to answer r to the query p+q=?' is not true of me: it is either false or lacks a truth value altogether. But, we may reasonably

ask, why should I have to get the sum right for numbers that are *too large for my mind to grasp* to succeed in meaning *addition* by '+'? This requirement places *too great an epistemic burden* on the speaker – indeed one that is *impossible* to meet. Given that no human can meet this requirement, no human can succeed in meaning *addition* by '+', or in understanding the term, according to [KD<sup>plus</sup>]. Thus, by characterizing the dispositional theory in a way that leaves this requirement in place, Kripke creates a *straw man*.

If we leave the requirement in place, we can try to meet it by adopting a proposal like that offered by Simon Blackburn, which characterizes the answer I *would* give for indefinitely large sums as the answer I would give by 'reiterat[ing] procedures I am disposed to use a number of times'. [1984 p. 289].

I think a simpler solution, though, is simply to restrict the biconditional in [KD<sup>plus</sup>] for terms like '+' whose extensions are determined by a computational or algorithmic procedure.

S's meaning *addition* by '+' consists in the fact that **for any p, q, r that S is capable (in principle) of grasping**, S is disposed<sup>57</sup> to answer r to the query 'p+q=?' iff r is the *sum* of p and q, and *addition*<sup>58</sup> is the simplest function corresponding to these behavioral limitations.

What this version of the dispositional theory entails for '+' is that I need not possess dispositions that get indefinitely large summation queries right – queries concerning numbers that are too large for even the fastest computer to handle -- to succeed in

---

<sup>57</sup> As we will see, I will soon need to introduce a ceteris paribus clause into this equivalence, not to surmount the finitude objection, but the error objection.

<sup>58</sup> As I noted in chapter 1.1, Kripke is not raising a doubt about the extension of '+', about, say, whether 125 is indeed the arithmetical sum of 68 and 57. He is raising a doubt about what he terms the 'metalinguistic' sense of correctness: What fact about me makes it the case that I used '+' to denote the addition and not the quadition function? [p. 8]

meaning *addition* by '+', a demand which no human can meet, but only an *initial segment* of the function involving the longest computations that my mind is (in principle) capable of grasping. Thus reformulated, the dispositional theory would be invulnerable to the finitude objection.

So it seems, but Paul Boghossian, among others, has offered an example which purports to show that the finitude objection applies to non-mathematical terms, like 'horse', whose extensions are not determined by a computational or algorithmic procedure. Thus, he says: 'if it is indeed the property *horse* that I am disposed to apply the term to, then I should be disposed to apply it to *all* horses, including horses so far away and so far in the past that it would be nonsense to suppose I could ever get into causal contact with them. Otherwise, what is to say that my disposition is not a disposition to apply the term to the property *nearby horse*, or some such? But no one can have a disposition to call all horses 'horse', for no one can have a disposition with respect to inaccessible objects.' [1989 p. 528-9]

In order to determine whether Kripke's formulation of the dispositional theory is vulnerable to this objection, we first need to determine what a more general version of the theory is supposed to look like. Here we run into some difficulties, because Kripke never describes the more general version.

In his discussion of Kripke's dispositional theory, Paul Horwich offers the following schematic generalization of '[KD<sup>plus</sup>']:

[KD] S's meaning F by w consists in the fact that S is disposed to apply w to x  
 iff x is an f. (where 'f' is to be replaced by a predicate (e.g. 'dog') and 'F'  
 by a name of the meaning of that predicate (e.g. 'DOG'))

He then argues [1998, 1995, 1990] that [KD] is implausible and should be rejected because it presupposes an *inflationary* theory of truth. According to Horwich [e.g. 1995 p. 363-364], [KD] is designed to provide a way of 'reading off' a predicate's extension from whatever property constitutes its meaning. In other words, [KD] implies that we can determine the extension of w from how S is disposed to apply w:

w is true of x iff S is disposed to apply w to x

This implies that there are non-semantic necessary and sufficient conditions for being true of – some account of the form

w is true of precisely the f's iff R(w,f)

where R(w,f) is deducible from whatever non-semantic property constitutes the meaning of w. But the existence of any such theory is plausible, Horwich maintains, only if *being true of* is a substantive relation, one that has some non-semantic underlying nature, some naturalistic analysis. And that, he maintains, is precisely what deflationism, which he upholds, denies.

We can use some terminology developed by Crispin Wright to see Horwich's point more vividly. Consider the following statement:

- (i) For any act  $x$ ,  $x$  is pious iff  $x$  is loved by the Gods.

In *Truth and Objectivity* (Appendix to Ch. 3: "The Euthyphro Contrast: Order of Determination and Response-Dependence", pp. 108ff), Wright points out that (i) is common ground between Socrates and Euthyphro. So what is the difference between them? Developing an idea deployed by Mark Johnston<sup>59</sup>, Wright proposes that the difference lies in the fact that Socrates accords a certain *order of priority to the left hand side* of the biconditional, which he terms '**detectivism**' about piety, while Euthyphro accords a certain *order of priority to the right hand side*, which he terms '**projectivism**' about piety.

Johnston<sup>60</sup> used this distinction to differentiate primary and secondary qualities. Consider, e.g., the following statements:

- (ii)  $x$  is square iff  $x$  would look square to standard observers under standard conditions.
- (iii)  $x$  is red iff  $x$  would look red to standard observers under standard conditions.

---

<sup>59</sup> See his 'Objectivity Refigured: Pragmatism without Verificationism' in J. Haldane and C. Wright (eds) *Reality: Representation and Projection* (New York: Oxford University Press, 1992).

Johnston proposed that (ii) should be read in a *detectivist* direction involving left to right priority, and (iii) in a *projectivist* direction involving right to left priority. The latter involves ‘writing human responses into the account of a substantial truth condition’. [Truth and Objectivity, p. 109] In other words, judgments about color are claimed by Johnston to be **response-dependent**:<sup>61</sup> the verdicts of standard observers under standard conditions **determine the extensions** of our color terms. On this reading, (iii) is true by definition: the biconditional stands for conceptual equivalence. Shape terms, on the other hand, according to Johnston, are response-independent: the verdicts of standard observers under standard conditions **track** an independently constituted extension. So (ii) is a contingent truth: the verdicts of standard observers under standard conditions *may* fail to track an independently constituted extension.

Now, consider again Kripke’s version of the dispositional theory.

[KD]        S’s meaning F by w consists in the fact that  
                   x is an f iff S is disposed to apply w to x.

In light of the terminology developed above, we can say that what Horwich finds problematic in [KD] is that it makes the meaning of our terms response-dependent, since their extensions are defined by how agents (it doesn’t matter if we replace S with a community of agents C) are disposed to apply them, and this, he maintains, entails an

---

<sup>60</sup> Wright notes that Johnston developed this view in a seminar delivered at Princeton in 1986, but has since moved away from it.

<sup>61</sup> This term is Wright’s. I am using it interchangeably with what I referred to as the projectivist view above, without assuming any of the additional properties -- in particular those involving necessity and apriority, which I find highly problematic -- that Wright builds into it in *Truth and Objectivity*.

inflationary theory of truth. It will not help, Horwich says, to introduce a *ceteris paribus* clause into [KD] (as we find in ii and iii above), for, he agrees with Kripke and, as we will later see, Boghossian, that such a clause cannot be described *non-circularly*. [See 1995 p. 361, 1998 p. 219]<sup>62</sup>

According to Horwich [1998 p. 219], Kripke's imposition of the '**reading off**' **requirement** causes him to assume that the right dispositional analyses have the form [KD]. Recall that, as Kripke defines this requirement, the value that I (or my linguistic community) means '+' to have for m and n is, *by definition*, the answer I (or they) would give if queried about m and n. [WRPL p. 25] Horwich believes that we can in fact give a dispositional account of meaning, and that Kripke's argument to the contrary is fatally flawed because it wrongly presupposes that a dispositional theory must meet the reading-off requirement. What we need to do, according to Horwich, is reject this unreasonable constraint.

I agree with Horwich that [KD] is implausible and should be rejected.<sup>63</sup> However, unlike Horwich, I do not think the fatal flaw lies in Kripke's 'reading off' requirement. In other words, it is not the *projectivist* reading of the biconditional in [KD] or its communal version that I find problematic; indeed, in what follows (§§2.2 – 2.4), I defend a version of the dispositional theory that defines the *extensions* of our terms in terms of communal use (and hence that leaves the reading off requirement in place); the only exception being terms whose extensions are determined through computational or

---

<sup>62</sup> Horwich proposes that we replace the crude version of the dispositional theory [KD], which has the abstract form S means F by w iff R(S,f,w) with the sequence: S means *plus* by w iff R<sub>1</sub>(S,w), S means *chair* by w iff R<sub>2</sub>(S,w), etc., where the right hand side of the biconditional makes no reference to f, but only to non-relational regularities in the speaker's use of w. [See, e.g., 1998 p. 219]

<sup>63</sup> Thanks to Paul Horwich for helping me see that a dispositional account of meaning does not have to take the form [KD].

algorithmic procedures. The account I offer invokes a *ceteris paribus* clause, which I believe *can* be specified non-circularly.

One can, along with Horwich, interpret the biconditional in [KD] or its communal version in a *projectivist* fashion, and perhaps this is how Kripke intended it to be read. However, one might also interpret it in a *detectivist* fashion, and it is this latter reading of [KD] or its communal version that I find problematic. For one thing, the standards for what counts as a correct application of a term, on this reading, would have to issue their requirements independently and in advance of communal verdicts for an open-ended range of situations. But how can they ‘reach ahead of us’, so to speak, and determine of themselves their every actual and counterfactual application? And how can we account for our ability to be appropriately sensitive to the specific requirements that they demand?<sup>64</sup>

I believe a more reasonable and defensible version of the dispositional theory is the following:

[D<sub>CP</sub>] S’s meaning by *w* what is meant by it in his linguistic community C consists in the fact that (1) in *ceteris paribus* conditions, S is disposed to apply *w* in ways that agree with how his linguistic community C is disposed to apply *w*, and (2) C’s meaning *F* by *w* consists in the fact, in *ceteris paribus* conditions, C is disposed to apply *w* to *x* iff *x* is an *f*.<sup>65</sup>

---

<sup>64</sup> At least two (metaphysical and epistemological) difficulties emphasized by Wright [See, e.g., 2001, which collects many relevant articles] in his critique of detectivism, though he does not apply these objections to Kripke’s discussion of the dispositional theory as I do here.

<sup>65</sup> Thanks to Paul Horwich for pointing out that I am committed to the additional thesis (2) and need to add it to [D<sub>CP</sub>]. Applied, e.g., to the word ‘chair’, [D<sub>CP</sub>] would read as follows:

On this social theory of meaning, the norms that govern linguistic practice are determined by the community. Language mastery is *parasitic* on communal practice, on more or less uniform patterns of collective behavior. No appeal to mysterious Platonic entities – ‘extensions’ having a life of their own, so to speak -- is necessary (as the detectivist assumes) once we interpret the extensions of expressions as *projections* of communal use.

Several points about [D<sub>CP</sub>] require further clarification:

1) I am employing the notion of a ‘**community**’ loosely in [D<sub>CP</sub>]. It is an elastic notion, like, and not in any significant respect different from, Wittgenstein’s concept of a *language-game*. Sometimes ‘community’ refers to S’s linguistic community *as a whole*; sometimes to a *proper subclass* of his community (e.g., biologists, mathematicians, historians, art critics, etc.). For many technical terms, this subclass involves a class of specialists – people who have expert knowledge of the term.<sup>66</sup> For terms where it makes no sense to speak of someone having ‘expert knowledge’ of them – e.g., ‘chair’, ‘table’,

---

S’s meaning by w what is meant by it in his linguistic community C consists in the fact that, in *ceteris paribus* conditions, S is disposed to apply ‘chair’ in ways that agree with how his linguistic community C is disposed to apply ‘chair’, and C is disposed (in *ceteris paribus* conditions) to apply ‘chair’ to x iff x is a chair.

It is important to note, however, that I intend the biconditional in (2) to be interpreted in a *projectivist* fashion involving left to right priority, and not in a detectivist fashion, which I find implausible. It should also be noted that (i) ‘f’ is a member of the extension of w, (ii) ‘F’ is not a name for the reference of w, but for the sense of w; more specifically, it is a name for the meaning w has *in S’s linguistic community*, and (iii) [D<sub>CP</sub>] is a general theory of meaning that applies to any word, not just those having a reference. So a less misleading way of expressing (2) is: C’s meaning F by w consists in the fact that C would judge that w satisfies use-context f in *ceteris paribus* conditions iff w satisfies use-context f.

<sup>66</sup>The idea of turning to the verdicts of ‘experts’ as our criterion of correctness includes that of deferring to the instruments they use in arriving at their results: computing machines, telescopes, microscopes, seismometers, or whatever the case might be.

‘beautiful’ – ‘community’ is to be interpreted as referring to the subject’s linguistic community *as a whole*. ‘Community’ can also refer to a narrow use-context involving two members (or ‘players’, as Wittgenstein would call them). How small can this use-context be? Can S himself constitute a ‘community’? If so, must his language be accessible to others? These questions form the substance of Wittgenstein’s famous private language argument. Though they are interesting and important, I do not think it would be appropriate, nor do I have the space, to deal with them here. What I would like to note is that [D<sub>CP</sub>] does not rule out the possibility of a private language, IF we permit an individual to constitute a ‘community’. Of course, the possibility of creating standards in this way assumes that the members of the ‘community’ have a propensity to *agree* in judgment; in absence of such agreement, standards could not emerge.<sup>67</sup>

2) It is common for theories of meaning to take the form ‘S’s meaning F by w, where F is **the** meaning of w, consists in the fact that ...’ I find these theories extremely confusing, because it does not make sense to speak of ‘the’ meaning of an expression, but only of its meaning *relative* to some context of usage. That is why I formulate [D<sub>CP</sub>] as ‘S’s meaning by w what is meant by it in his linguistic community C consists in the fact that ...’, or, equivalently, ‘S’s meaning F by w, where F is the meaning w has in S’s linguistic community C, consists in the fact that ...’ Words can have a multitude of meanings. For example, ‘major’ can mean ‘high officer in an army’, ‘field of study’, ‘greater in

---

<sup>67</sup> As Wright puts it, ‘It is, for epistemological purposes, a *basic* fact about us that ordinary forms of explanation and training do succeed in perpetuating practices of various kinds – that there is a shared uptake, a disposition to concur in novel judgments involving the concepts in question. The mythology of the “rules as rails” attempts an explanation of this fact. But the truth is the other way round: it is the basic disposition to agreement which sustains all rules and rule-governed institutions. The requirements which our rules impose on us would not be violated if there were not this basic agreement; they would not so much as *exist*.’ [2007 p. 7]

importance, rank, or stature’, ‘requiring greater attention or concern; serious’, ‘more inclusive in scope’, ‘of greater number, quantity, or extent’, ‘a musical mode where ...’, ‘someone who has reached full legal age’, etc. To borrow Wittgenstein’s game analogy, it does not make sense to speak of a ‘correct move’, but only of a correct move relative to some game. A word does not carry the same meaning into every game; it is the language game in which the word is used that fixes its meaning; it gets its meaning by how the players *use* it. Unlike [D<sub>CP</sub>], [KD] gives no clue as to how the standards for what counts as a correct application of a term is determined.<sup>68</sup>

3) On the view I will develop in §2.4 – which I term ‘**constructive dispositionalism**’ -- it is communal practices that bring meanings *into existence*. Meanings are progressively created in a variety of ways, and in a variety of contexts. And they may *change* over time. They are not found ready-made, and they are not fixed for all time. As I stated in §1.3.2, meanings are not stagnant, like dead matter, but alter and adapt to their environment with the passing of time, like **living organisms**. Those that fail to do so simply cease to exist, or exist only in the space of conceptual possibilities.<sup>69</sup> This analogy has important implications, as we will see in §2.4, for the ‘underlying nature’ of meaning-constituting facts. And it is another reason I find the detectivist reading of the communal version of Kripke’s dispositional theory incoherent. Only where the extension

---

<sup>68</sup> If the biconditional in the communal version of [KD] is to be interpreted along detectivist lines, it leaves us (to borrow a phrase from Crispin Wright, though he himself does not apply it in this context) with ‘no accountable idea of what would constitute the direction taken by the rule off its own bat, as it were... once the direction is conceived, after the fashion of Platonism, as determined autonomously, and our performance, whether communal or solitary, as merely an exercise in tracking’. [2001, p. 161]

<sup>69</sup> The origin of ‘democracy’, or Nietzsche’s discussion of the origin of ‘good’ and ‘bad’, which I discussed in §1.3.2, illustrates this point well.

of a word can be determined by an *algorithmic or computational procedure* does it make sense to say that the standards for what counts as a correct application of it are determined in advance and may come apart from communal use; in all other cases, the standards are *parasitic* on communal use. An individual can only use language meaningfully when his linguistic behavior is held accountable to the norms set by the appropriate members of his linguistic community. How the community is disposed to apply a term ‘supplies the essential background against which it makes sense to speak of an agent’s application of a term as correct or incorrect’, for words of his public language, and ‘for the community itself there is no authority, so no standard to meet’.<sup>70</sup> The community forms the bedrock, the final court of appeal on whether an application of a term is correct.

4) Of course, a community may misapply a term under certain conditions. In the next section (where I address Kripke’s error objection) I will show how [D<sub>CP</sub>] can accommodate error. Briefly, the problem that error poses for [D<sub>CP</sub>] is this. According to [D<sub>CP</sub>], my applications of a term are correct if they agree with those of (the appropriate members of) my linguistic community – since how they are disposed to apply a term serves as the standard of correctness. But intuitively, we can conceive of circumstances in which my linguistic community is bound to make the same mistakes I do. It too, for example, might be disposed to call both horses and deceptively horsy-looking cows on dark nights ‘horse’. But how can we make sense of the community going collectively off track if their use serves as the standard of correctness? If the community is disposed to apply ‘horse’ to horsy-looking cows on a dark night, then it appears we must conclude

---

<sup>70</sup> Wright’s expressions [1980 p. 219-20].

that that simply *is* a correct application of the term, since how they are disposed to apply a term determines what counts as a correct application of it. So, if we are to make sense of the notion of communal error, it looks like we need to introduce something stronger than communal consensus into a dispositional theory of meaning to serve as the standard of correctness. This poses a problem for [D<sub>CP</sub>], for, if the standards of correctness are not constituted by our reaction to the case, then it looks like they must be constituted independently of our reaction, and this, it appears, would commit me to the kind of platonist metaphysics that I want to steer away from in [D<sub>CP</sub>]. I will turn to this problem in the next two sections (2.2 and 2.3), where I address the error objection in greater detail.

5) [D<sub>CP</sub>] accommodates the notion of *degrees* of understanding: it doesn't rule out the intuitive thought that understanding is not an all or nothing matter, but may be *partial*. Suppose, for example, that I do not apply the term 'mammal' to whales because I am ignorant of the fact that whales are mammals – I think instead that they are fish -- but in most other cases I apply the term 'mammal' as those who have expert knowledge of the term do (here, marine biologists). As [D<sub>CP</sub>] is currently formulated, it follows that I do not *understand* the term 'mammal', since the *right-to-left* direction of the biconditional in [D<sub>CP</sub>] breaks down. Cases involving ignorance can also be used to show that the *left-to-right* direction of the biconditional in [D<sub>CP</sub>] breaks down. I would fail, for example, to understand the terms 'gold', 'whale', or 'water', if I repeatedly apply these terms to objects that experts would exclude from their extension – e.g. fool's gold, creatures having the gestalt of a whale but different internal characteristics, a substance having the observable features of water but a different chemical composition, etc. But what we

should take these examples to show is that I don't *fully* understand the terms in question. We non-experts do not have knowledge of the criteria specialists employ in applying many technical terms that have high traffic in our language; our standard regimen of training does not bestow it on us. Yet this does not preclude us from regularly and systematically applying them in agreement with one another, or in effectively communicating. This is because we do not need complete understanding in order to communicate effectively, nor does [D<sub>CP</sub>] require any such thing.<sup>71</sup>

Now, how does [D<sub>CP</sub>] fare with Boghossian's extension of the finitude objection? I stated that it is only terms whose extensions can be determined by a computational or algorithmic procedure that can be thought of as response-independent; it is only here where it makes sense to say that communal use may fail to track an independently constituted extension. Boghossian appears to be assuming that the extension of 'horse' is determined in advance of communal verdicts; that it somehow takes care of itself. But, as I have argued here, whether an object falls under the extension of 'horse' is a decision that is made by the relevant experts on the matter (here, biologists), and not something that is determined in advance. And so long as my use conforms with that of the experts, I mean what they do by it. So [D<sub>CP</sub>] is not vulnerable to this objection.

---

<sup>71</sup> [D<sub>CP</sub>] also brings out a sense in which [SKEP] differs from the more general version of the skeptical problem [SKEP\*]: 'Is there any fact about me that my meaning *something* by "+" consists in?'. A positive answer to [SKEP\*] need not entail a positive answer to [SKEP]. Suppose, for example, that the body of mathematicians is disposed to call P a proof, whereas I am inclined not to. Then, according to [D<sub>CP</sub>], I would not mean *proof* by 'proof', since it's not the case that I am disposed to apply the term 'proof' to a string of sentences iff the experts would. But I do mean *something* by 'proof' (assuming I am not applying the expression in any old way).

## 2.2 OVERCOMING THE ERROR OBJECTION

### 2.2.1 The Error Argument

The error objection purports to show that the dispositional theory cannot make sense of the notion of a mistaken application: it renders error impossible by failing to allow that what a person means by a term and how he is in practice disposed to use that term may diverge. Kripke formulates this objection as follows:

Most of us have dispositions to make mistakes. For example, when asked to add certain numbers some people forget to ‘carry’. They are thus disposed, for these numbers, to give an answer differing from the usual addition table. Normally, we say that such people have made a mistake. That means, that for them as for us, ‘+’ means addition, but for certain numbers they are not disposed to give the answers they *should* give, if they are to accord with the table of the function they actually *meant*. But the dispositionalist cannot say this. According to him, the function someone means is to be *read off* from his dispositions; it cannot be presupposed in advance which function is meant. In the present instance a certain unique function (call it ‘skaddition’) corresponds in its table exactly to the subject’s dispositions. (Waive the difficulty that the subject’s dispositions are finite: suppose he has a disposition to respond to any pair of arguments.) So, where common sense holds that the subject means the same addition function as everyone else but systematically makes computational mistakes, the dispositionalist seems forced to hold that the subject makes no computational mistakes, but means a non-standard function (‘skaddition’) by ‘+’. [p. 29-30]

This objection has been generalized by Paul Boghossian [1989 p. 531] to other terms, e.g., ‘horse’. Suppose I mean *horse* by ‘horse’. Then, presumably I have a disposition to call horses ‘horse’. But it will also be true that there are certain circumstances (e.g. sufficiently dark nights) and certain objects (e.g. sufficiently horsy looking cows) such that I am disposed, under those circumstances, to call those objects ‘horse’ too.

Intuitively, this is a disposition, under those circumstances, to make a *mistake*. But, according to Boghossian, Kripke's dispositionalist cannot call that a 'mistaken' application. Rather, he must conclude that I mean something non-standard – e.g. *horse or cow*. Thus, where common sense holds that I mean *horse* by 'horse' but make mistakes when applying the term under certain conditions, the dispositionalist is forced to conclude that I make no mistakes, but am correctly following a 'bent rule'.<sup>72</sup>

The error argument has the same basic form as the finitude argument, but, unlike the finitude argument, which shows that the right-to-left direction of the biconditional in [KD] fails, the error argument purports to show that *both directions* of the biconditional in [KD] fail, i.e. that:

$$\forall o, S \text{ is disposed to apply } w \text{ to } o \leftrightarrow o \text{ is an } f$$

$$\dots \rightarrow \dots$$

Once again, the argument takes the form of a *reductio*.

### *Error argument against the dispositional theory*

---

<sup>72</sup> This expression is Simon Blackburn's.

The error objection has been applied more generally to other disposition-like theories. For example, Alex Miller points out [1998 p. 192] that it applies to *causal theories* of meaning of the form:

S means X by 'Y' iff it is a *law* that Y's *cause* tokenings of 'X' for S.

The main problem for this theory, according to Miller, is that it leaves no room for the notion of a 'mistaken' application of 'Y'. But there are going to be certain circumstances -- e.g. sufficiently dark nights -- and certain objects -- e.g. sufficiently horsy looking cows -- such that I invariably token 'horse' on encountering them. It will then follow from the theory, Miller contends, that by 'horse' I mean *horse or horsy-looking cow*. Fodor, in *A Theory of Content and Other Essays* (Cambridge: MIT Press, 1990), refers

- i Suppose I mean F by w, where w is a term I am disposed to misapply under certain conditions (e.g. ‘horse’ to horsy-looking cows on dark nights).
- ii According to the dispositional theory [KD], I mean F by w if and only if I am disposed to apply w to all and only f’s.
- iii Since I am disposed, under certain conditions, to apply w to non-f’s, and to fail to apply w to f’s, I do not, according to [KD], mean F by w, or understand it.
- iv But this contradicts the supposition (i) that I do mean F by w. So the dispositional theory [KD] must be wrong.

### 2.2.2 Critique of the Error Argument

There are two ways a dispositionalist can respond to the error objection. He can either accept the consequence that a subject who makes mistakes -- such as forgetting to carry -- under certain conditions, does not mean *addition* by ‘+’, but something non-standard. Or he can stick with common sense and say that a subject can mean *addition* by ‘+’, even though his dispositions to calculate with ‘+’ may not, under certain conditions, correspond perfectly to the addition function. In the latter case, according to Kripke, he would have to give up [KD<sup>plus</sup>] as a way of formulating his theory. So how can he formulate it? In the last section, I proposed that we reformulate the dispositional theory as a social theory of meaning:

---

to this as the **disjunction problem**, and considers it to be the main problem facing *naturalistic* accounts of meaning.

[D<sub>CP</sub>] S's meaning by w what is meant by it in his linguistic community C consists in the fact that (1) in *ceteris paribus* conditions, S is disposed to apply w in ways that agree with how his linguistic community C is disposed to apply w, and (2) C's meaning F by w consists in the fact, in *ceteris paribus* conditions, C is disposed to apply w to x iff x is an f

This reformulation of the dispositional theory blocks the error objection, so long as we leave the *ceteris paribus* clause in place. For, if we remove it from [D<sub>CP</sub>], then the error objection would surface at both the individual and communal levels as follows:

(1) I might mean M by w (where M is the meaning my linguistic community assigns to w) and yet be disposed, under certain conditions, to apply w in ways *not shared* by my community. I might, for example, along with my linguistic community, mean *addition* by '+', or *red* by 'red', and yet, when intoxicated, be disposed to fail to carry digits that end in 7's, or see certain shades of red as green under a certain type of lighting, when in those same conditions, my linguistic community would carry, or see those shades as red. It appears to follow from [D] (i.e., [D<sub>CP</sub>] without the CP clause), however, that I do not mean *addition* by '+', or *red* by 'red', if I am disposed, under certain conditions, to apply these terms in ways that fail to agree with how my linguistic community applies them. Rather, I mean something non-standard. So it appears that [D], like [KD], cannot make sense of the notion of a mistake at the individual level.

(2) Moreover, generalizing Kripke's error objection, it seems that [D] cannot make sense of the notion of a mistake at the communal level. According to [D], my applications of a term are correct if they agree with those of my linguistic community – since my linguistic community's applications of a term serve as the standard of correctness. But intuitively, we can conceive of circumstances in which my linguistic community is bound to make the same mistakes I do. It too, for example, might be disposed to call both horses and deceptively horsy-looking cows on dark nights 'horse'. As Paul Boghossian remarks, 'If I can be taken in by a deceptively horsy-looking cow on a dark night, what is to prevent 17,000 people just like me from being taken in by the same, admittedly effective, imposter? ... (This is presumably what makes magicians possible.)' [1989 p. 536] But how can we make sense of the community going collectively off track if their use serves as the standard of correctness?<sup>73</sup> If the community is disposed to apply 'horse' to horsy-looking cows on a dark night, then it appears we must conclude that that simply *is* a correct application of the term, since how they are disposed to apply a term determines what counts as a correct application of it. So, if we are to make sense of the notion of mistake at the communal level, it looks like we need something *stronger* than communal agreement to serve as a standard of correctness. So it looks like, as in (my interpretation of) [KD], we need to introduce some kind of platonist (communal-transcendent) standard of correctness back into the dispositionalist account of meaning, and interpret the biconditional in the communal version of [KD] along detectivist lines.

---

<sup>73</sup> This problem for communal versions of the dispositional theory has also been raised by Blackburn [1984a].

(2) presents us with a dilemma, which Crispin Wright [2007] describes as follows. One horn consists in the ‘**communitarian response**’, according to which the standards of correctness have ‘somehow to be located *within* the propensities for assessment of the case possessed by others in my language community ... But the evident awkwardness with this idea is that it seems to reduce the correctness of an assessment to a kind of marching in step, and to exclude all room for the idea of judgments that are true anyway, no matter what we come to say about a question, or whether we even consider it at all’ [2007 p. 5] The other horn consists in the ‘**platonist response**’ which says that ‘even (hypothetically) shared assessments are constitutively quite independent of the requirements they concern – that even in the far-fetched scenario where a whole speech community assents to a particular utterance, and where everybody is clear-headed, attentive, and generally competent, the communal impression of what ought to be said is one thing and what really ought to be said is something else’. [2007 p. 6] The problem is, if, as I maintain, the standards of correctness are *not* constituted (as the platonist thinks) independently of our reaction to the case, then what can be available to constitute them *except* our reaction?

By introducing a *ceteris paribus* clause into [D], as in [D<sub>CP</sub>], I submit that we can surmount this dilemma.<sup>74</sup> To see how, consider a substance that possesses a dispositional characteristic: salt. Salt has a disposition to dissolve in water. But it will not dissolve in a sample of water that is saturated with, say, benzene or petrol. So the counterfactual ‘If salt is placed in water, it will dissolve’ does not always hold. But we don’t, in light of this fact, *reclassify* salt as a substance that does not possess the disposition to dissolve in

---

<sup>74</sup>Wright’s own solution is detailed in [2001 Ch. 7].

water. Rather, we say that something intervened (benzene or petrol or whatever the case may be) that blocked the manifestation of salt's disposition to dissolve in water. Likewise, we wouldn't reclassify a drug regarded as a stimulant simply because a subject who absorbed it failed to exhibit stimulated behavior, if in the circumstances the subject also absorbed a depressant or received some lethal shock.<sup>75</sup> The general point is that an object could possess a dispositional characteristic, and yet the manifestation of that characteristic be blocked or inhibited by some feature of the object's situation (for example, benzene or petrol in the case of salt's disposition to dissolve in water).

The main problem with the error objection is that it underestimates the complexity of a disposition. Though it may be true that a person who means *horse* by 'horse' may be disposed to apply 'horse' to horsy-looking cows on a dark night, it is also true that he would refrain from applying 'horse' to those *same* objects in other situations – e.g., on a sunny day, when they are in close range, in full view, etc. Likewise, though a person who means *addition* by '+' may fail to carry digits that end in 7's when intoxicated, it is also true that he would perform those *same* calculations correctly when he is not intoxicated, well rested, free of distractions, etc. In light of these facts, the dispositionalist can argue that how someone is disposed to apply a word (what dispositional property he possesses), and hence what he means by it, is not determined by how he is disposed to use it under conditions that cause him to err, but by how he is disposed to apply it in *ceteris paribus* conditions. These will include, in the case of non-observational terms like '+', conditions where: the subject is not tired, intoxicated, or distracted (by, e.g., surrounding noise), has not suffered a cognitive injury, and has been given adequate time to reflect on his

---

<sup>75</sup> These examples are offered by Graham Forbes [1984, p. 229].

application of the term (in the case, on his calculation). In the case of observational terms like ‘red’<sup>76</sup> we need to add conditions where: the subject is not visually impaired (i.e., he possesses normal visual equipment), is observing the object attentively, and where the object observed is in full view, in good lighting conditions (like those that actually typically obtain out-of-doors and out-of-shadow), relatively stationary, and not too far away. For other terms, we will need to add other conditions. CP is meant exclude all conditions *where a subject would (sincerely) judge that a word satisfies a given use-context (that, e.g., it applies to a certain object), yet fail to endorse his application under further investigation.* The conditions offered in CP are not meant to be exhaustive; empirical research will reveal what further conditions need to be added. It is not possible to foresee *all* ways in which people can err, so we cannot give a complete characterization of CP. To take an example, suppose the scientific community believes it has discovered a new planet in a distant region of space from satellite images of that region. They agree to call it a ‘planet’ and give it a name. But, as they later find out, the lens of the camera used to take the satellite images was not producing accurate images, and caused certain stars to look like planets. The scientific community made a mistake, i.e., it applied ‘planet’ to an object, yet failed to endorse its application upon further investigation of the use-context (this context includes the images and instruments used to derive them). So we need to add another condition to CP: ‘instruments used to determine whether a word satisfies a given use-context are in proper working order’. The point is, it is difficult to predict all conditions that need to be added to CP in advance.

---

<sup>76</sup> I draw some of these conditions from Crispin Wright’s discussion of the ‘judgment-dependent’ nature of color terms [1989b]

The error argument is flawed because it draws the following invalid inference (in premise iii))<sup>77</sup>:

- i. Under certain aberrant conditions, a subject has a disposition to make a mistake (e.g., to call cows ‘horses’).
- ii. Therefore, he has a disposition to make *that* mistake.

Whether a subject has a disposition to make a mistake is not (in general) determined by how he is disposed to apply a term in aberrant conditions, but by how he is disposed to apply the term in CP conditions (roughly, the normal ones); just as, in the case of the salt crystal, whether it has a disposition to dissolve in water is determined by how it is disposed to behave in the normal conditions of water. These are the conditions under which, according to the canons of science, dispositions are ascribed to objects.<sup>78</sup>

Returning to our earlier example, though it may be true that a person who means *horse* by ‘horse’ may be disposed to apply ‘horse’ to horsy-looking cows on a dark night, it is also true that he would not apply ‘horse’ to those *same* objects in CP conditions. So

---

<sup>77</sup> Many discussions of the error argument draw the same *invalid* inference. It is drawn, for example, by Boghossian [1989 p. 534ff].

<sup>78</sup> A referee from *Mind* pointed out the following objection to me: ‘If tiny robots were made that looked and acted like houseflies, I would (under normal conditions) be disposed to apply the term “fly” to them, and while I might be able to tell the difference if I put one under a microscope, that condition does not seem to be “normal”.’ In reply, I would like to stress that [D<sub>CP</sub>] is offered as a theory of meaning and understanding in the *actual* world, not some possible futuristic world. As I develop the theory more fully in chapter 2.4, meanings are constructed over time, and they may change over time, and with it the standards for what counts as a correct application. They are not fixed in advance for all time or across possible worlds. In the author’s possible world, which houses robot flies, human beings might have telescopic vision, that penetrates the interior makeup of objects in their visual field, so that they can tell the difference between flies and robot flies (those might be the ‘normal’ conditions). In any event, whether I understand the term ‘fly’ in that world depends on whether I use it as an expert would, and they may very well decide to include robot flies in the extension of ‘fly’. Given that the jury is out on this case, we cannot use it as a counterexample to [D<sub>CP</sub>].

we would not be entitled to conclude that he has a disposition call those objects ‘cows’, and, in turn, that he means something non-standard by ‘horse’, as premise (iii) does. Similar remarks apply to someone who means *addition* by ‘+’ but miscalculates under certain conditions. If a subject is disposed to calculate correctly in CP conditions, then his making systematic mistakes in aberrant ones does not warrant our *reclassifying* him as a ‘quadder’ (short for someone who is disposed to follow some non-standard rule). Kripke’s error objection fails because it draws this unwarranted inference (in premise (iii)).<sup>79</sup>

This observation can also be used to deflect the second objection -- that we cannot make sense of the notion of a mistake at the communal level. Even though we can conceive of circumstances in which my linguistic community is disposed to make the same ‘mistakes’ I do (e.g. to apply ‘horse’ to deceptively horsy-looking cows on dark nights), it does not follow that they are *generally* disposed to make these ‘mistakes’. For it is also true that they would refrain from applying ‘horse’ to those *same* objects in other situations – e.g., on a sunny day, when they are in close range, in full view, etc. And so we can take what my community means by a term to be determined by how they are disposed to apply it in CP conditions. The possibility of error arises when the CP conditions have not been satisfied, not when communal verdicts are out of line with independently constituted extensions -- an idea that I find implausible.

---

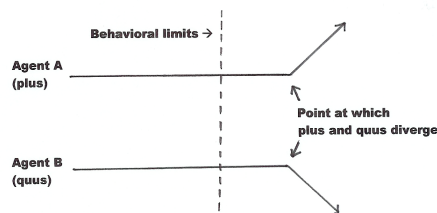
<sup>79</sup> On a charitable reading, we might say Kripke does not draw this invalid inference in the error argument, for the mistakes are supposed to be, as he says, ‘systematic’. I suppose by this he means that our calculator is supposed to be analogous to one who suffers from, say, dyslexia. Suppose this person, in no matter what context of usage, is prone to write a 9 instead of a 6, so that he always ends up with the wrong result. Are we not warranted in concluding that he has a disposition to make a mistake (i.e., to infer ii from i)? Perhaps, but in any event, if the CP conditions, as I have characterized them, are satisfied, he will not have the disposition to make that mistake, for one of the conditions in CP is that he is not suffering from a brain defect. So this particular calculator does not pose a threat to [D<sub>CP</sub>].

According to the error objection, the dispositionalist cannot distinguish a subject A who means *addition* by ‘+’ but makes mistakes from a subject B who means some nonstandard function that accords with the ‘mistakes’ (where A and B could be the same person).<sup>80</sup> On the strategy I have proposed, the dispositionalist can frame the relevant distinction as follows. He can argue that, in mastering or acquiring the plus rule, A has acquired a disposition of a particular sort. But his having acquired and continuing to possess the disposition does not mean that A will manifest it on *every* occasion in which its manifestation is called for: on some occasions, the manifestation of the disposition may be blocked or inhibited. This, the dispositionalist can argue, should not cause us to reclassify A as a quadder, since in CP conditions, A would give answers that accord with the plus rule. B, on the other hand, given that he means something non-standard -- let us call it quus -- would, in such conditions, give answers that accord with the quus rule.

This strategy, of course, grants to the skeptic that B *can* mean something non-standard even if he receives the same training as A. But one may refuse to grant the

---

<sup>80</sup> The claim here is that what one means by a term does not supervene on how he is disposed to use it because one’s dispositions would have been *exactly the same* in another possible world even if one meant something different. In the case where summation queries are within my cognitive grasp, Kripke motivates this objection through the possibility of error. But he also considers the following scenario for summation queries beyond my cognitive grasp. Let ‘quaddition’ be redefined so as to be a function that agrees with addition for all pairs of numbers small enough for a subject to have any disposition to add them, and let it diverge from addition thereafter (say, it is 5). Then, he says, it is possible for an agent B to mean *quaddition* by ‘+’ in another possible world while agent A, who has *exactly the same* dispositions as B in this world, means *addition* by ‘+’. The situation is depicted in the figure below.



This scenario, however, is hardly plausible. How is an agent able to form the *intention* that numbers behave in a quus-like way at some point well beyond his cognitive grasp, which presumably he can’t even identify?

skeptic even this much, for it flies in the face of some commonly held beliefs; namely, that:

- (i) Behavioral dispositions are holistically connected, and
- (ii) Mutual pressures toward conformity within a linguistic community help bring one's dispositions in line with correct patterns of usage, so long as he is possessed of normal cognitive powers.

Suppose a skeptic comes along and asks: 'How do you know that B, who is disposed to use "+" *exactly the same* way as A (an adder) has not taken "+" to mean the bent rule that any number added to a million sums up to five?' Well, what else is true of how B is disposed to behave, if he interprets '+' in this silly way? Suppose he has a million dollars in his bank account. Will he not be disposed to avoid adding more, for fear of going bankrupt? Will he not be disposed to view winners of the multi-million-dollar jackpot as the biggest losers? Etc. His world would be a very strange world indeed – in fact, a deranged world! Described in the abstract, a person's disposition to apply '+' in a non-standard way may seem reasonably contained<sup>81</sup>, but when we examine how this disposition is likely to affect other dispositions, it becomes apparent that the assumption that a bent interpretation can be contained without ramification is illusory. Bent interpretations of '+' are bound to affect the rest of a person's verbal and non-verbal behavior in bizarre ways; indeed in ways that would cause us to doubt his sanity if his

---

<sup>81</sup> How a subject is disposed to apply '+' cannot be so contained that it has no role or function in his life, otherwise, along with Wittgenstein, I would say he attaches no meaning to '+'.

behavior departs radically from his community's. As Simon Blackburn [1984a p. 79] puts it: 'The concealed bent-rule follower is a theoretician's *fiction!*'

In so far as an individual is a member of a linguistic community, communal practices kick in to help bring his erroneous dispositions in line with the community's. Communal practices include tendencies to correct, criticize, and adjust deviations. They serve as 'normative model[s], the function of which is to constrain or regulate behavior'.<sup>82</sup> So in a community, a deviant who interprets '+' in a bent way, or applies 'horse' to cows, or 'cat' to fish, is criticized. In being corrected by others, he acquires new beliefs, and modulo these beliefs, how he is disposed to apply a term comes further in line with communal practices. Mutual pressures toward conformity, consisting of reinforcements and punishments, help bring one's erroneous dispositions in line with correct patterns of usage.<sup>83</sup>

In light of (i) and (ii), it is hard to see how someone can mean a bent rule by '+' and yet be disposed to apply '+' in *exactly the same* way as an adder. The skeptic cannot simply help himself to this assumption. It needs an argument, and so far as I can tell, there is none in the offing, for any such argument would need to establish (at least) the falsity of (i) and (ii) – i.e., it would need to convince us that, for an individual with normal cognitive powers, communal sanctions can fail to induce a process of adjustments in his behavior involving '+' that leads to convergence with communal practices -- so as to make sense of the idea that he can be merely disposed to 'go astray at the bend'.<sup>84</sup>

Kripke himself admits to being puzzled by how an intelligent creature could get the

---

<sup>82</sup> Barbiero's expression [\*\*\*\*].

<sup>83</sup> For further discussion, see Simon Blackburn [1984b] and, for a lengthier account of how this process work, Barbiero [\*\*\*\*].

very training we have for the addition function, and yet grasp the appropriate function in a quus-like way.

Of course we can define the quus function, introduce a symbol for it, and follow the appropriate rule for computing its values. I have done so in this very essay. What it may be unintelligible to us is how an intelligent creature could get the very training we have for the addition function, and yet grasp the appropriate function in a quus-like way. If such a possibility were really completely intelligible to us, would we find it so inevitable to apply the plus function as we do? Yet this inevitability is an essential part of W's own solution to the problem.

The point is even stronger with respect to a term like 'green'. Can we grasp how someone could be presented with a number of green objects, and be told to apply the term 'green' just to 'things like these', and yet apply the term learnt as if it meant 'grue'? It would seem that if we find our own continuation to be inevitable, in some sense we cannot. [p. 98n]

So he too gives us reason to doubt the requisite tasks can be executed by the skeptic.

## 2.3 OVERCOMING THE CIRCULARITY OBJECTION

### 2.3.1 The Circularity Argument

After concluding that  $[KD^{plus}]$  fails, Kripke remarks that it will not help the dispositional theorist to introduce a *ceteris paribus* clause into  $[KD^{plus}]$  that excludes conditions that can cause a subject to err, for the resulting analysis would turn out to be circular.

Once again the difficulty cannot be surmounted by a *ceteris paribus* clause, by a clause excluding 'noise', or by a distinction between 'competence' and 'performance'. No doubt a disposition to give the true sum in response to each

---

<sup>84</sup> Blackburn's expression [1984b].

addition problem is part of my ‘competence’, if by this we mean simply that such an answer accords with the rule I intended, or if we mean that, if all my dispositions to make mistakes were removed, I would give the correct answer. (Again I waive the finiteness of my capacity.) But a disposition to make a mistake is simply a disposition to *give an answer other than the one that accords with the function I meant*. To presuppose this concept in the present discussion is of course viciously circular. If I meant addition, my ‘erroneous’ actual disposition is to be ignored; if I meant skaddition, it should not be. Nothing in the notion of my ‘competence’ as thus defined can possibly tell me which alternative to adopt. Alternatively, we might try to specify the ‘noise’ to be ignored without presupposing a prior notion of which function is meant. A little experimentation will reveal the futility of such an effort. Recall that the subject has a *systematic* disposition to forget to carry in certain circumstances: he tends to give a uniformly erroneous answer when well rested, in a pleasant environment free of clutter, etc. One cannot repair matters by urging that the subject would eventually respond with the right answer after correction by others. First, there are uneducable subjects who will persist in their error even after persistent correction. Second, what is meant by ‘correction by others’? If it means rejection by others of ‘wrong’ answers (answers that do not accord with the rule the speaker means) and the suggestion of the right answer (the answer that does accord), then again the account is circular. ... [p. 29-32]

Recall that the first adequacy condition on any fact that solves [SKEP] is that it be a state of an agent, which figures in an *analysis* of meaning addition by ‘+’. At least one major obstacle facing anyone attempting to offer an analysis is that of **circularity**. The goal in giving an analysis is to provide illuminating conceptual equivalences, and illumination will depend on the analysans’ utilizing concepts which, in the best case, are in some way *prior to* and *independent* of the notion being analysed – or, if that’s too much to ask for, then some form of explication which does not take one straight back to the notion being analysed.<sup>85</sup> We would, of course, be taken straight back to the notion being analysed if it reappears *explicitly* in the analysans, as in: ‘A book is pornographic iff it contains

---

<sup>85</sup> As Crispin Wright notes in ‘Truth: A Traditional debate Reviewed’ in S. Blackburn and K. Simons (eds) *Truth* [1999]. Examples of analyses include attempts to define mental expressions in terms of expressions referring to actual or possible behavior (*logical behaviorism*), or in terms of neurobiology; attempts to define the concepts of classical mathematics in terms of logic and set theory (*logicism*); and attempts to define statements about the empirical world in terms of immediate experience, actual or possible (*phenomenalism*)

pornography'; 'An animal is human iff it has human parents'. If the dispositional theorist is to avoid the charge of circularity, then he must not characterize the *ceteris paribus* conditions in a way that makes explicit use of the term being defined. This much is incontestable. For example, the *ceteris paribus* conditions cannot be characterized as 'conditions where we have been provided with sufficient mental capacities to allow us to *add* any two numbers'.

An analysis would also be circular or unilluminating if the term being defined appears *implicitly* in the analysis, as in: 'X is good if and only if ideal people like X.' (Although the word 'good' does not recur explicitly, it seems hidden in the notion of an 'ideal' (= 'maximally good') person). 'Art is that which artists produce.' (What is an artist? Someone who produces *art*. So the definition takes us back to the question: What is art?). 'Art is that which hangs in galleries.' (What is a gallery? A place which houses *art*. Again, we are taken back to the question: What is art?). Likewise, Kripke argues that characterizing the *ceteris paribus* conditions as 'conditions when a subject is corrected by others' issues in a circular analysis; presumably because 'answer given when corrected by others' *implicitly contains* the notion 'answer conforming to the meaning of "+"

Kripke maintains that a dispositional theorist cannot characterize the *ceteris paribus* conditions in a way that avoids the charge of circularity. We can sum up his argument as follows:

*Circularity argument against the dispositional theory*

- i. Any fact that constitutes my meaning *addition* by '+' must be characterized by a non-circular description.

- ii. Facts about how I am disposed to apply ‘+’ in ceteris paribus conditions cannot be characterized non-circularly.
- iii. Therefore facts about how I am disposed to apply ‘+’ in ceteris paribus conditions cannot constitute my meaning *addition* by ‘+’.

### 2.3.2 Critique of the Circularity Argument

It is far from obvious that the skeptic’s demand for a non-circular characterization of the ceteris paribus conditions cannot be met. Many mistakes we make in applying terms arise because of the presence of aberrant conditions that have a *generalizable and predictable* effect on creatures with similar cognitive endowments. Empirical research can help us determine what those conditions are. We can then rule them out of our characterization of the ceteris paribus conditions [CP]. For example (as I suggested in §2.3), in the case of *non-observational terms* like ‘+’, we can take CP to include conditions where: the subject is not tired, intoxicated, or distracted (by, e.g., surrounding noise), has not suffered a cognitive injury, and has been given adequate time to reflect on his application of the term (in this case, on his calculation). In the case of observational terms like ‘red’<sup>86</sup> we need to add conditions where: the subject is not visually impaired (i.e., he possesses normal visual equipment), is observing the object attentively, and where the object observed is in full view, in good lighting conditions (like those that actually typically obtain out-of-doors and out-of-shadow), relatively stationary, and not

---

<sup>86</sup> I draw some of these conditions from Crispin Wright’s discussion of the ‘judgment-dependent’ nature of color terms [1989b]

too far away. As mentioned earlier, this list of conditions is not meant to be exhaustive. Empirical research will determine what further conditions need to be added to CP. In answer to Kripke's challenge, then, a dispositionalist can analyze a subject's meaning *addition* by '+' as follows:

[D<sub>CP</sub><sup>plus</sup>] S's meaning *addition* by '+' consists in the fact that **for any p, q, r that S is capable (in principle) of grasping**, S is disposed (in ceteris paribus conditions) to answer r to the query 'p+q=?' iff r is the *sum* of p and q, and *addition* is the simplest function corresponding to these behavioral limitations.

The ceteris paribus conditions as described above neither explicitly nor implicitly contain the term being defined, and they are not trivially specified, so there is no reason to think the account is circular.<sup>87</sup>

We can only speculate about how the skeptic would respond to this analysis, since he has not provided a *general* reason for adopting premise (ii) – he has only given us reason to reject *some* ways of filling in the ceteris paribus conditions (as, e.g., 'conditions where my brain is stuffed with extra matter that enables me to add', 'conditions where I am corrected by others') -- and in the absence of such an argument, what reason do we have

---

<sup>87</sup> The skeptic might object that the analysis I have offered is not successful unless 'facts' about how I am disposed to apply a term are defined **reductively**, i.e., 'vertically', so to speak, from complex to simpler levels, and ultimately in terms of the vocabulary of physics. But this would constitute another *unwarranted* assumption on what form a meaning-constituting 'fact' must take. One can offer an analysis of a concept, and in particular of what meaning something by an expression consists in, by disclosing the 'lateral' relations it shares with other concepts. There is nothing that forces an analysis of meaning to be reductive. For discussion of this unwarranted reducibility assumption, see Wright [2001 p. 146ff]

to accept premise (ii), or to reject  $[D_{CP}^{plus}]$ ? Surely, the *onus is on the skeptic* to offer an argument.<sup>88</sup>

### *Boghossian's Argument*

Here, Paul Boghossian [1989], in his critique of Kripke's argument against the dispositional theory, attempts to fill in the lacuna. He presents an argument based on the *holistic* nature of belief-fixation that 'there could not be naturalistically specifiable conditions under which a subject will be disposed to apply an expression only to what it means; and, hence, that no attempt at satisfying such conditions can hope to succeed.' [1989 p. 539] His argument runs as follows. Consider Neil and a particular expression, say 'magpie', in Neil's mental repertoire. According to Boghossian, a dispositionalist needs to show that the following equation holds: Neil means *magpie* by 'magpie' iff in optimal conditions, Neil is disposed to token the belief that x is a magpie iff x is a magpie. Boghossian argues that the dispositionalist cannot succeed at this. His argument is based on the observation that the interpretation of a speaker's beliefs is a *holistic* affair. With suitably mistaken background beliefs present, a speaker will be inclined to apply particular expressions incorrectly. As he puts it:

---

<sup>88</sup> Paul Horwich has pointed out the following potential problem to me. If we define a 'correct' application of a word to be one which agrees with how one's linguistic community is disposed to apply the word in *ceteris paribus* conditions, then, we cannot in turn define 'ceteris paribus conditions' to be conditions where one's linguistic community applies words *correctly*, otherwise our account will be circular. We can, however, overcome this difficulty by filling in the *ceteris paribus* conditions in the way I have above.

Under normal circumstances, belief fixation is typically mediated by background theory -- what contents a thinker is prepared to judge will depend upon what other contents he is prepared to judge. ... Thus, Neil may come to believe 'Lo, a magpie' as a result of seeing a currawong, because of his further belief that that is just what magpies look like; or because of his belief that the only birds in the immediate vicinity are magpies; or because of his belief that whatever the Pope says goes and his belief that the Pope says that this presented currawong is a magpie. And so on. The thought that something is a magpie can get triggered by a currawong in any of an indefinite number of ways, corresponding to the potentially indefinite number of background beliefs which could mediate the transition. [1989 p. 539-40]

These background beliefs must, then, be excluded from the conditions that are supposed to yield correct usage of an expression. But, Boghossian argues, a dispositionalist cannot do this without violating the non-circularity constraint. Following is my summary of his argument. [1989 p. 539ff]

*Boghossian's argument against the dispositional theory*

- i A dispositional theorist has to specify, without the use of semantic or intentional materials, a property M such that possession of M is necessary and sufficient for being a disposition to apply an expression in accord with its correctness conditions; for example, a property M such that a subject will be disposed to think 'Lo, a magpie' only in respect of magpies.
- ii But the holistic character of belief formation implies that a subject will be disposed to think 'Lo, a magpie' in respect of an *indefinite* number of non-magpies, provided only that appropriate background beliefs are present.
- iii Since there looks to be a potential infinity of background clusters of beliefs which could potentially mediate the transition from non-magpie to magpie beliefs, specifying an optimality condition for 'magpie' will involve specifying

a situation in which it is guaranteed that none of this potential infinity of background clusters of belief is present.

- iv But the dispositional theorist cannot succeed at this task without invoking intentional materials. So a reductive theory of meaning cannot take the form of a dispositional theory.

### *Critique of Boghossian's Argument*

Again the crucial question to ask is: Why should we think that the *ceteris paribus* conditions cannot be non-circularly specified? Boghossian's thought appears to be that, since there is a potential infinity of possible background beliefs that can cause a subject to err (to misapply a term), there is no finite way to specify the *ceteris paribus conditions* that avoids the use of intentional material.

The main problem with this contention is that it overlooks two important assumptions that are granted by the skeptic. Recall the skeptical problem is: What constitutes my mastery of any word of my language? According to [D<sub>CP</sub>], *what* I mean by a word is determined by how I am disposed to apply it in *ceteris paribus* conditions; my *mastery* of a word consists in the fact that I am disposed to apply it in *ceteris paribus* conditions in agreement with how the relevant body of my linguistic community is disposed to apply it; and to learn the meaning of a word is to acquire such dispositions.<sup>89</sup> According to

---

<sup>89</sup> How a subject is disposed to apply a word may be explained in terms of his *beliefs* about it, but it is not necessary for a dispositionalist to make reference to them in an analysis of what a subject means by a term: he need only refer to how he is disposed to use the term.

Boghossian, on the other hand, my mastery of a word (say, ‘magpie’) does not consist in this because there is a ‘potential infinity of background clusters of beliefs’ that can cause me to misapply it, and I can’t rule these out in a non-question-begging way.

But it is crucial to ask: What *kinds* of beliefs does Boghossian have in mind here? It cannot be beliefs that *clash* with the assumption that I have mastered the word in question. Examples of such beliefs for ‘cat’ are: ‘Cats live in water’, ‘Cats have fins’, ‘Cats are blue’. I will refer to these as *standing beliefs*, since they (erroneously) form part of my concept of a cat. It cannot be beliefs that clash with the assumption that I have mastered the word ‘cat’ that are supposed to be excluded by the *ceteris paribus* conditions, because the skeptic is *giving* me the assumption that I have mastered the word in question (that I understand it), and his query is: What does my mastery (understanding) of the word consist in? He is, in particular, giving me the assumption that I have mastered the word ‘+’, and he is asking what my mastery of ‘+’ consists in.

It is, on the other hand, *not* part of my mastery of the term ‘cat’ that I be able to distinguish cats from small poodles in a distance on a cloudy day, or from robot cats, etc. For it is conceivable that I have mastered the term yet *regularly* misapply it under certain conditions (e.g. in cloudy weather, on dark nights, to imposters, etc.). The error objection capitalizes on this fact. Certain conditions can generate beliefs -- I will refer to these as *circumstantial* beliefs, since they are not standing, but are rather based on the circumstances I find myself in -- about an object that interfere with my correctly applying a term to that object. (For example, on a cloudy day, I might form the *circumstantial* belief that it is correct to apply ‘horse’ to the cows in the distance, but I don’t have the *standing* belief that it is correct to apply ‘horse’ to cows.) But if the mistakes are *regular*

(as opposed to being, say, ‘arbitrary stabs in the dark’), then through empirical research, we can arrive at a suitable *generalization/law* about the conditions under which they are made. And on the basis of this law, we can make *predictions* about when they are likely to be made again. We can then exclude these conditions in our characterization of the CP conditions.

We might interpret Boghossian as claiming that we cannot formulate such a law because, for one thing, there is no *finite* way to specify it. But why should we assume this? If we cannot foresee all the conditions in which a subject is likely to misapply a word, it doesn’t follow that there are infinitely many such conditions. We might also interpret Boghossian as saying that, even if the conditions *can* be finitely specified, they cannot be *non-circularly* specified. But we only have reason to believe this if, as I argued above, we assume that the subject has not mastered the term in question -- that he has a number of erroneous *standing* beliefs about the object in question that must be ruled out of our characterization of the CP conditions. For example, if a subject has the erroneous belief that ‘cat’ applies to objects that live in water, have fins, and are blue, then, Boghossian appears to be saying, the CP conditions would have to include the clause that ‘S does not believe cats live in the water, have fins, and are blue’. But the skeptic is giving us the assumption that the subject *has* mastered the term in question, so there is no reason to think we must rule out standing beliefs of this kind.<sup>90</sup>

---

<sup>90</sup> There are ways other than that of demanding an analysis that we can interpret the question of what ‘constitutes’ what we mean by our words in [SKEP], so that even the skeptic’s first adequacy condition is not forced on us. Some interpret the skeptical problem as raising the challenge of whether there are any facts that meaning facts **supervene** on.

The basic idea behind a supervenience thesis is that, although we talk in terms of two levels of facts or properties (‘lower level’ and ‘upper level’), one of them is fundamental in that once it is fixed then so is the other. In other words, if two things have the same basic or ‘lower level’ properties, then they must also have the same ‘upper level’ properties. Or: It is not possible that two things should have the same lower-level properties without also having the same upper-level properties.

For example, it is claimed that consciousness supervenes on complex neurophysiological facts; a melody supervenes on a sequence of notes; dispositions of a thing supervene on its physical constitution; the fact that a computer is running a particular program supervenes upon microphysical facts about its components; a painting's representational properties supervene on the geometrical arrangement of light-reflecting surfaces; etc. (Note: The converse of the supervenience relation does not necessarily hold: two things may be identical in the upper-level respect without being identical 'all the way down'. For example, a computer program might be 'variably realized': two computers might be running the same program functionally speaking, but with rather different configurations of hardware.)

Supervenience is held to be a weaker relation than definitional (reductive or non-reductive) analysis on the ground that it offers the possibility that a property may depend on other properties, without being *definable* in terms of those properties.

Some critics of the skeptical argument (e.g. Horwich [1998], Soames [1997], Blackburn [1984b]) have defended a supervenience-based version of the dispositional theory as a straight solution, which may be expressed as follows:

[D<sub>s</sub>] It is not possible that two people should be identical in terms of how they are disposed to apply a term, and yet mean something different by it.

There is a passage where Kripke's skeptic has been interpreted as arguing against the plausibility of [D<sub>s</sub>]:

I might have introduced '\*' to mean nothing in particular even though the answer I arbitrarily choose for 'm\*n' is, through some quirk in my brain structure, uniquely determined independently of the time and other circumstances when I am asked the question. It might, in addition, even be the case that I consciously resolve, once I have chosen a particular answer to 'm\*n', to stick to it if the query is repeated for any particular case, yet nevertheless I think of '\*' as meaning no function in particular. What I will not say is that my particular answer is 'right' or 'wrong' in terms of the *meaning* I assigned to '\*', as I will for '+', since there is no such meaning. [p. 24 n]

The skeptic's contention in the described scenario appears to be that my responses to '\*'-queries are not real answers but only verbal reflexes; they are just sounds/utterances, capable neither of correctness nor incorrectness, since no assertion/judgment is made. Scott Soames [1997] tries to make sense of this contention by inviting us to imagine, as a result of some special brain surgery, being programmed to verbally respond to the query 'What is x \* y?' by uttering a definite numeral z for each pair of numerals x and y used in the question. Despite this, it might still be the case, he maintains, 'that I don't *mean* anything by "\*"': my responses may be nothing more than verbal reflexes. I might notice that I always give determinate responses to queries using "\*". But I don't regard these as correct or incorrect – since I don't intend to assign any meaning to "\*". [1997 p. 222]

The skeptic's argument, thus interpreted, is hardly compelling. Suppose a scientist has programmed me to give the sum of x and y every time I have to compute their sum. Then I see no reason why it should not be said of me that I mean *addition* by '+'. It should not matter that the ability to produce these sums was *instantaneously* acquired through implantation as opposed to *over time* through my interaction with a community. The skeptic owes us an explanation for why this should make a difference.

The explanation, following Martin and Heil [1998] and Dummett [1991], might run as follows. In using a word with a certain meaning, an agent must not only be using it in 'conformity with a rule', but his behavior must be 'based on the rule', he must 'act on the rule'. On this reading, the skeptic reasons that since the agent's behavior in the described scenario is merely 'conforming to a rule' as opposed to being 'based on a rule', he is not using '+' with a certain meaning.

The distinction between 'conforming to a rule' and 'acting on a rule' has surfaced in many domains – it plays, e.g., a central role in Kantian ethics, in determining whether an act has moral worth. But what does '**acting on a rule**' or '**basing your behavior on a rule**' amount to for the skeptic? Having a rule *guide* you? If so, then, as I show in the next section, the skeptic's argument fails: in an important sense, our rule-following behavior is not 'based on a rule' at all.

## 2.4 OVERCOMING THE JUSTIFICATION OBJECTION

### 2.4.1 The Justification Argument

The most powerful objection that Kripke's skeptic raises against the dispositional theory is the justification objection. Unlike the finitude and error objections, which try to show that my dispositions to calculate with '+' do not correspond perfectly to the addition function, the justification objection maintains that even if my dispositions *do* correspond perfectly to the addition function, that would not be enough to show that I mean *addition* by '+'. [p. 57] The skeptic insists that any fact that 'constitutes' my meaning one function, rather than another, by a given function sign, be such that, whatever answer I am in fact disposed to give, it should 'tell me' what answer I *ought* to give, and facts about how I am disposed to apply a term (even if in accord with its correctness conditions) do not, he insists, have this property.

The key passages where Kripke discusses the justification objection are the following:

To a good extent this reply [the dispositional analysis] immediately ought to appear to be misdirected, off target. For the skeptic created an air of puzzlement as to my *justification* for responding '125' rather than '5' to the addition problem as queried. He thinks my response is no better than a stab in the dark. Does the suggested reply advance matters? How does it *justify* my choice of '125'? What it says is: "125' is the response you are disposed to give, and (perhaps the reply adds) it would also have been your response in the past.' Well and good, I know that '125' is the response I am disposed to give (I am actually giving it!), and maybe it is helpful to be told – as a matter of brute fact – that I would have given the same response in the past. How does any of this indicate that – now or in the

past – ‘125’ was an answer *justified in terms of instructions I gave myself*, rather than a mere jack-in-the-box unjustified and arbitrary response? [p. 23]

So it does seem that a dispositional account misconceives the skeptic’s problem – to find a past fact that *justifies* my present response. As a candidate for a ‘fact’ that determines what I mean, it fails to satisfy the basic condition on such a candidate, stressed above on pg. 11, that it should *tell me* what I ought to do in each new instance. [p. 24]

As I noted in §1.1, the skeptic interprets the justification constraint (that any candidate meaning-constituting fact F must satisfy) in a strong internalist<sup>91</sup> sense:

[J<sub>I</sub>] F justifies my applying word w thus-and-so, given what I mean by it, if and only if (1) I have *direct (non-inferential) access* to F, and (2) I can *deduce a priori* from F that I ought to apply w thus-and-so for an open-ended range of situations.

Note that (2)  $\not\Rightarrow$  (1). Suppose, for example, F is a simple rule for multiplying numbers in exponential form. If I had direct (non-inferential) access to F, I would be able to deduce a priori from F that I ought to give answer z to any query ‘ $a^x \times a^y = a^?$ ’, where x and y are numbers within my cognitive grasp. It doesn’t follow that I *do* in fact have direct access to F. Note also that (1)  $\not\Rightarrow$  (2). I might, e.g., say that to mean *green* by ‘green’ is to bring up a color image of green whenever I use the word ‘green’. The image would be readily available or transparent to me, but I cannot deduce a priori from it that I ought to apply ‘green’ to all and only green objects, since it can be interpreted in a grue-like way.

As I noted in §1.1, [J<sub>I</sub>] is composed of two sub-conditions: a *metaphysical* constraint (concerning the *normativity* of meaning) which requires that F have a certain **nature**, and an *epistemological* constraint which requires that F have a certain **role**.

**Metaphysical constraint** (concerning the *normativity* of meaning): F contains a norm, a standard of correctness, which determines in advance how w is to be applied in an open-ended range of situations.

**Epistemological constraint:** I have direct access to F (so that it can guide future applications of w).

If F contains instructions on how to apply w in an open-ended range of situations (metaphysical constraint), and I have direct access to F (epistemological constraint), then I will be able to deduce a priori from F that I ought to apply w thus-and-so in any given situation, as [J<sub>I</sub>] requires.

Condition (1) in [J<sub>I</sub>] (i.e., the epistemological constraint) insists that any purported meaning-constituting fact F be **readily available** to me, i.e., that I have **unmediated knowledge** of it, such as I do when I experience a pain, a tickle, a red after-image, or ringing in the ears. The subject has direct (non-inferential) access to such states because, ‘since they are events in his consciousness, he is in the nature of the case conscious of them’.<sup>92</sup> That this constraint be met by F is suggested in the quote from p. 24, where Kripke says that ‘it should **tell me** what I ought to do in each new instance’. Further evidence for this condition can be found on p. 40:

---

<sup>91</sup> For an alternative interpretation of the justification constraint, see Michael Levin [2007].

<sup>92</sup> As Crispin Wright observes [2001 p. 149].

The idea that we lack '**direct**' **access** to the facts whether we mean plus or quus is bizarre in any case. Do I not know directly, and with a fair degree of certainty, that I mean plus? Recall that a fact as to what I mean now is supposed to justify my future actions, to make them *inevitable* if I wish to use words with the same meaning with which I used them before. This was our **fundamental requirement** on a fact as to what I meant.

That the skeptic insists, in addition, that any purported meaning-constituting satisfy the metaphysical constraint is suggested in the quote from p. 11, where in talking about 'directions I gave myself', he says these directions 'must be **contained** in any candidate for the fact as to what I meant'. This requirement is also suggested in other places, where the skeptic insists that any putative meaning-constituting fact 'compel', 'establish' [p. 13, 21], 'show' [p. 41], 'entail' [p. 53], 'mandate' [p. 57], 'determine in advance' [p. 65], 'guarantee' [p. 95] that I mean *addition* by '+'.

The fact F that we are looking for must, on this interpretation of the justification constraint, be such that it provides the basis for an a priori deductive inference to claims about what I mean, and be transparent to me. Under these conditions, F will give me a **reason** for thinking '125' is the answer I ought to give, as the skeptic demands. [p. 11] I will then be able to cite F as a *justification* for the answer I give in order to show it isn't just a shot in the dark, 'a mere jack-in-the-box unjustified and arbitrary response'. [p. 23] This is why Kripke sometimes [see e.g. pp. 11-13] expresses the skeptical problem in *epistemological* terms: *How do I know* that I now mean by '+' the same thing that I meant in the past by '+'? In answering this question I offer a *reason*: 'I know *because ...*'

The justification constraint can be read as being even stronger than [J<sub>1</sub>]. Kripke says that the basic condition on any candidate for a fact that determines what I mean is that 'it

should *tell* me what I ought to do in each new instance'. [p. 24] This may require us to add a third condition to [J<sub>I</sub>]:

[J<sub>I+</sub>] F justifies my applying word *w* thus-and-so, given what I mean by it, if and only if (1) I have **direct (non-inferential) access** to F, (2) I can **deduce a priori** from F that I ought to apply *w* thus-and-so for an open-ended range of situations, and (3) F **guides** my applications of *w* in each new case, so that my applications are based on *inferences* from F.

[J<sub>I+</sub>] is stronger than [J<sub>I</sub>] in the following sense. [J<sub>I</sub>] demands a *reason* for my applying *w* thus-and-so. A **reason** is something I might produce to show that my application of a word is justified if someone were, say, to throw down a skeptical challenge. By 'reason' the skeptic has in mind 'instructions' [p. 23] or 'directions' [p. 11] I gave myself that 'tell me' [p. 24] how I should apply *w*. A **guide**, on the other hand, is something I consult with *each* fresh application of a term, as I might consult a recipe in baking a cake, or a dictionary in translating another language. A reason R (set of instructions or directions that determine how I should apply *w*) can serve as a guide in this sense, but it need not: I might not consult R every time I apply a term, even though I might produce it to justify any given application. A guide, unlike a reason, also need not be expressed in words – for example, a guide might be a traffic-light, a sign-post, or a map. But the skeptic does not appear to have this sense of a guide in mind: the constraint on the meaning-constituting fact is that it should 'tell me' what I ought to do in each new instance.

With [J<sub>I</sub>] in place, we can summarize the skeptic's justification argument against the dispositional theory as follows:

*Justification argument against the dispositional theory*

- i My dispositions to answer questions 'What is  $m + n$ ?' constitute my meaning *addition* by '+' only if I have direct (non-inferential) access to facts about how I am disposed to answer such questions, and I can deduce a priori from such facts that I ought to give such-and-such an answer to such questions.
- ii Having direct (non-inferential) access to facts about how I am disposed to answer addition questions does not provide the basis for an a priori deductive inference for the answer I ought to give to such questions.
- iii Therefore, my dispositions to answer questions 'What is  $m + n$ ?' do not constitute my meaning *addition* by '+'.<sup>93</sup>

Scott Soames [1997] tries to motivate the skeptic's reasoning behind (ii) as follows:

Suppose I had amnesia and didn't remember what I meant by '+' in the past. When asked, at present, to calculate '68 + 57' I might feel inclined to give the answer '125'. Still, I might wonder whether this was justified, in terms of what I meant by '+' in the past. Would it be sufficient to be told that this was the answer I was disposed to give in the past? It doesn't seem that it would, for **I might wonder whether in the past this was one of those calculations on which I was disposed to make mistakes. A similar observation holds regarding the point of view of someone viewing me from the outside and wondering whether my**

---

<sup>93</sup> Replacing [J<sub>I</sub>] with [J<sub>I+</sub>] in the justification argument gives rise to a stronger version of the justification argument. Unlike premise (ii) in the [J<sub>I</sub>]-version, which raises the question: 'Can you deduce (a priori) how you ought to apply a term from knowledge of the meaning-constituting fact?', premise (ii) in the [J<sub>I+</sub>]-version raises the additional question: 'Do you in fact proceed this way?'. The skeptic concludes that the answer is no. This much seems correct. When I apply terms, there is nothing that appears to *guide* me, nothing that I do in fact consult, with each fresh application of a term, that *tells me* how to go on. I am, so to speak, on my own.

**present answer of ‘125’ is correct.** Simply being given a specification of the relevant past dispositions is not enough to allow such an observer to conclude that my present answer accords with my past meaning or intentions. [1997 p. 222]

This way of motivating (ii) rests on the *objection based on error*. But what if, as I pointed out in §2.3, we can avoid the discrepancies pointed out in the objection based on error by identifying conditions under which how I am disposed to use a word determines what I mean by it. In that case, the argument above would fail, if premise (ii) was motivated *merely* by the error objection. But the justification argument is supposed to be *stronger* than the argument based on error: it is supposed to show that even if the answers I am disposed to give *do* correspond perfectly to the addition function (in, say, *ceteris paribus* conditions), that would still not be enough to show that I mean *addition* by ‘+’. Thus, it cannot be motivated *merely* by the error objection, which can be overturned if we can specify conditions under which my dispositions to calculate with ‘+’ do correspond perfectly to the addition function.

The skeptic must, therefore, either have some other means of motivating premise (ii), or a stronger argument than that given above in mind if he wants to distinguish the justification objection from the error objection; something like the one below – the *ceteris paribus* version of the argument above.

*Ceteris paribus version of the justification argument*

- i My dispositions to answer questions ‘What is  $m + n$ ?’ in *ceteris paribus* conditions CP constitute my meaning *addition* by ‘+’ only if I have direct (non-inferential) access to facts about how I am disposed to answer such

questions in CP, and I can deduce a priori from such facts that I ought to give such-and-such an answer to such questions.

- ii Having direct (non-inferential) access to facts about how I am disposed to answer addition questions in CP does not provide the basis for an *a priori deductive inference* to what answer I ought to give to such questions.
- iii Therefore, my dispositions to answer questions ‘What is  $m + n$ ?’ in CP do not constitute my meaning *addition* by ‘+’.

#### 2.4.2 Critique of the Justification Argument

In assessing the ceteris paribus version of the justification argument, there are two questions we must ask:

- (1) (regarding premise ii) Is it true that facts about how I am disposed to apply a term in ceteris paribus conditions fail to satisfy [J<sub>I</sub>]?
  - (2) (regarding premise i) Is [J<sub>I</sub>] a reasonable constraint to place on any candidate meaning-constituting fact?

In what follows I will address each of these questions in turn.

#### *Evaluating premise (ii)*

In order to establish premise (ii), the skeptic must show that either:

- (a) I do not have direct (non-inferential) access to facts about how I am disposed to apply an expression in *ceteris paribus* conditions, or
- (b) I cannot deduce *a priori* from access to such facts that I ought to apply an expression thus-and-so for an open-ended range of situations, given what I mean by it.

Kripke advances some considerations in support of (a) when discussing the finitude objection.

I have heard it suggested that the trouble arises solely from too crude a notion of disposition: *ceteris paribus*, I surely will respond with the sum of any two numbers when queried. ... Perhaps, but how should we flesh out the *ceteris paribus* clause? Perhaps as something like: if my brain had been stuffed with sufficient extra matter to grasp large enough numbers, and if it were given enough capacity to perform such a large addition, and if my life (in a healthy state) were prolonged enough, then given an addition problem involving two large numbers, *m* and *n*, I would respond with their sum, and not with the result according to some quus-like rule. But how can we have any confidence of this? **How in the world can I tell what would happen if my brain were stuffed with extra brain matter, or if my life were prolonged by some magic elixir?** Surely such experiments should be left to science fiction writers and futurologists. We have no idea what the results of such experiments would be. They might lead me to go insane, even to behave according to a quus-like rule. The outcome really is obviously indeterminate, failing further specification of these magic mind-expanding processes; and even with such specifications, it is highly speculative. [p. 27]

This objection has a grip, however, only in the context of ‘magic mind-expanding’ thought experiments. I might not know how I would be disposed to apply a term if my

brain were stuffed with extra brain matter, or if my life were prolonged by some magic elixir. However, no magic mind-expanding powers entered into my characterization of the *ceteris paribus* conditions – conditions where a subject is not tired, distracted, or intoxicated, etc. It is through *empirical investigation* that we arrive at these conditions, not by consulting ‘science fiction writers and futurologists’. And it is reasonable to assume that I *do* know how I would behave in such conditions.

But Kripke appears to reject even the weaker claim that I have direct access to how I am disposed to apply words in conditions that don’t involve bizarre mind-expanding processes:

Am I supposed to justify my present belief that I meant addition, not quaddition, and hence should answer ‘125’, in terms of a **hypothesis** about my past dispositions? (Do I record and investigate the past physiology of my brain?) Why am I so sure that one particular hypothesis of this kind is correct, when all my past thoughts could be construed so that I meant plus or so that I meant quus? [p. 23]

Perhaps the thought here, following Zalabardo’s suggestion [1997 p. 479], is that if the answer I should now give were determined by my past dispositions, justifying my present response would require acquiring information about how I was disposed to answer in the past. But this information is not *directly* available to me now. Hence justifying my present response would involve forming a *hypothesis* about my past dispositions.

But would the skeptic maintain that I do not have direct (non-inferential) access to how I am *presently* disposed to apply a term? That is, after all, what is at issue in [SKEP]: we are looking for a fact about me *now* that constitutes my meaning what I do by my words. And surely I do have direct unmediated knowledge of how I am *currently*

disposed to apply words. This can be contrasted with the sort of knowledge I have of how others are disposed to apply words. Forming a belief about another person's linguistic behavior does involve forming a hypothesis (based perhaps on my knowledge of their past behavior, or on the assumption that the other person is very much like me) in a way that my belief about how *I* am disposed to apply words does not.<sup>94</sup> I have direct knowledge of how I am disposed to apply words, just as I have direct knowledge of my own conscious states and processes.<sup>95</sup>

The parenthetical remark in the quoted passage suggests, however, that Kripke thinks otherwise. His reasons, though not disclosed in the passage, can perhaps be reconstructed as follows. When we ascribe a dispositional property to an entity (e.g., brittleness to a pane of glass, solubility to salt), part of what we are saying is that if certain conditions are fulfilled, it will behave in a certain way (e.g., the pane of glass will break if struck with a moderate degree of force; salt will dissolve if placed in water). Like other empirical propositions, disposition-ascribing sentences appear to require a **truthmaker**: an event or state of affairs which is specified by the sentence and which, if it exists, makes the proposition true. But in the case of a disposition-ascribing sentence, the truthmaker cannot be an event or state of affairs in which the disposition is *manifested*; for it can be true that an entity possesses a property (brittleness or solubility) even though no manifestation of the disposition (breaking of the glass, dissolution of the salt crystals) ever has occurred or will occur. But what then is the truthmaker in the case of a disposition-ascribing sentence? According to David Armstrong [see, e.g., 1996], it is a

---

<sup>94</sup> As Carl Ginet observes [1992 p. 64], though not everyone agrees. Thus Crispin Wright writes: 'Subjects can generally know that they themselves possess [a dispositional psychological state] only in so far as and in the same way that others can know that they do'. [2001 p. 149]

‘categorical’ state of the entity to which the dispositional property is ascribed. Armstrong takes **categorical properties** of a concrete particular to include its shape, size, and material composition (e.g. nature and configuration of its molecules), which do not extend beyond what is actually the case or actually happening at some moment of time. These are contrasted, by Armstrong, with **dispositional properties**, which extend beyond what is actually happening or is actually the case at some moment of time, to what would happen or be the case if certain contingencies should arise in the future (or to what would have happened or been the case, if those contingencies had arisen in the past.).

Carl Ginet [1992] conjectures that underlying Kripke’s contention that beliefs about one’s past dispositions involve a *hypothesis* is the assumption that the truthmaker the dispositional account of meaning needs is a *categorical* state of the subject. With this assumption in place, Ginet [1994 p. 64] interprets Kripke as reasoning as follows toward (a):

- i. It is a necessary condition of the truth of the counterfactual conditional about me that there was some categorical basis for it. (That is, there were independent intrinsic properties of me that would causally explain why the condition or occurrence described in the antecedent would produce the occurrence described in the consequent, that would be sufficient for the truth of the counterfactual conditional.)
- ii If (i), then I could not know that the counterfactual conditional was true of me without knowing that there was some categorical basis for it.
- iii I could know that there was a categorical basis only as a hypothesis (that is, something confirmed for me by evidence as to what specific sorts of mechanism there are in me or by evidence as to how I and others like me have reacted in similar situations in the past).
- iv Therefore, I could know that the conditional was true of me only as an evidence-based hypothesis (and not in the sort of direct way I have memory knowledge of my recent past experience or have knowledge of my current intentions).

---

<sup>95</sup> Ibid.

Of course, against premise (ii), it can be argued (as Ginet points out) that it is not the case in general that one must know everything entailed by what one knows: for many instances where  $p$  entails  $q$ , one may know  $p$  without knowing  $q$  if one fails to know that  $p$  entails  $q$ . And against premise (i), it can be argued that the sort of counterfactual conditional that the dispositional account of meaning needs could be true without a categorical basis: the truthmaker could be, as U. T. Place argues for disposition-ascribing sentences more generally [see Armstrong 1992 p.7], a ‘counterfactual state of affairs’: the possession by the entity in question (the glass, say) of the dispositional property (being shattered when struck). For Place, the state of the entity whereby it possesses the dispositional property and the corresponding state of the microstructure of the entity are two distinct states of affairs. For Armstrong, by contrast, the state of affairs that makes a counterfactual true is a state of the microstructure of the entity that possesses the dispositional property: the dispositional property and the state of the microstructure are one and the same thing. Though Place’s thesis may be difficult to establish, clearly the burden of proof is on Kripke to establish the opposing (Armstrongian) view, i.e., to show that the truth of the counterfactual conditional needed by the dispositionalist account of meaning entails a categorical basis. In absence of such a proof, the above argument carries no weight. And so, to this point, the skeptic has failed to establish (a): that we do not have *direct access* to how we are disposed to apply a word in *ceteris paribus* conditions.

Jose Zalabardo [1997 p. 480] interprets the skeptic as attempting to establish (a) on the ground that we do not know how to identify the *ceteris paribus* conditions. This, of course, assumes that the ‘knowledge’ or ‘direct access’ operator in (a) has *wide scope*:

not only am I required to know (or have direct access to) *how* I am disposed to apply a term in ceteris paribus conditions, but I must also know *what* the ceteris paribus conditions are. But a dispositionalist need not insist on this: he can require that the subject only know how he is disposed to apply a word under ceteris paribus conditions, not in addition what the ceteris paribus conditions are. Nevertheless, it is likely that a subject *is* capable of identifying the conditions under which he is disposed to apply ‘+’ correctly. We are all prone to make occasional mistakes in calculations, and all we need to do in order to fill in the ceteris paribus conditions is call to mind past experiences where we or others made mistakes and a general characterization of what if anything those experiences had in common. With little reflection, we are likely to recognize that we tend to err in conditions where we are tired, distracted, intoxicated, or have been given inadequate time to make a calculation, among others. Again, it is only if we think of the ceteris paribus conditions as involving magic mind-expanding powers that we might be dumbstruck over how to characterize them. So, in light of all the foregoing considerations, it appears that we have not been given adequate reason to accept (a) (= I do not have direct (non-inferential) access to facts about how I am disposed to apply an expression in ceteris paribus conditions).

Turning now to (b) (= I cannot deduce *a priori* from access to such facts that I ought to apply an expression thus-and-so, given what I mean by it), suppose I have direct (non-inferential) knowledge of how I am disposed to apply a word. Can I *deduce a priori* from such knowledge that I ought to apply a word thus-and-so in any new case, given what I mean by it? According to the skeptic, I cannot. Dispositional facts, he reasons, do not determine how I *ought* to apply a word, given what I mean by it; they only determine

how I *would* apply the word. I take this to be the import of Kripke's observation that facts about how I am disposed to apply a word are **descriptive**, not **normative**. But though this may be the case for facts about how I am disposed to apply a word when conditions are not *ceteris paribus*, what reason do we have to believe that if: (i) How I would apply a word in CP conditions is how I should apply it, and (ii) I have direct access to how I would apply a word, nevertheless (iii) I cannot deduce a priori from it that I ought to apply the word thus-and-so?

Assuming we have successfully identified the *ceteris paribus* conditions, at least one reason the skeptic might offer is the following. [J<sub>1</sub>] requires that, for any word *w* whose sense I grasp, whatever fact *F* constitutes my grasping the sense of *w* behave like a *rule* that issues a determinate output for *any* given input, a rule that legislates all future uses, so that I can deduce a priori from knowledge of *F* that I *ought* to apply *w* thus-and-so, given what I mean by it, in any use-context. For any subject *S*, time *t*, and word *w*, we can describe the general form of this rule as follows:

$$\mathcal{R}_{(S,t,w)} = \text{Apply } w \text{ in context } c_i \text{ iff } c_i \text{ satisfies conditions } a_1 - a_n.$$

$\mathcal{R}_{(S,t,w)}$  is the 'direction' or 'instruction' I gave myself, that 'tells me' what to do in each new instance, as the skeptic demands. It is the 'reason' I offer for judging *w* to apply. By 'reason' the skeptic means an explicit formulation of the rule I am disposed to follow, which I can offer if someone demands a justification for my applying a word thus-and-so. It follows that  $\mathcal{R}_{(S,t,w)}$  determines a *complete* function whose domain contains all contexts

$c_i$  about which I can be asked: ‘Does  $w$  satisfy  $c_i$ ?’, and which assigns to each such context one of two values: ‘yes’ or ‘no’. Just as the rules  $f(x)=2x+1$ , or:

Rule: If neither King nor one of its Rooks has moved in the course of the game so far, and if the squares between them are unoccupied, and if neither the King nor any of those squares is in check to an opposing prince, then one may Castle.<sup>96</sup>

determine *complete* functions, so too, the skeptic reasons, for *every* word  $w$ , any fact that constitutes my grasping the sense of  $w$  is a rule-like entity that determines a complete function.

Now, we can describe facts about how a subject  $S$  is disposed to apply word  $w$  at time  $t$  in any context  $c_i$  as a partial function  $\mathbf{D}_{(S,t,w)}(c_i) = j$ , where:

$\text{Dom}(\mathbf{D}) = \{\text{all contexts of usage } c_i \text{ about which } S \text{ can be asked: ‘Does } w \text{ satisfy } c_i\text{?’}\}$

$\text{Ran}(\mathbf{D}) = \{\text{yes (‘1’), no (‘0’), don’t know (‘?’)}\}$

To illustrate, consider the predicate ‘red’. Let  $c_1, c_2, c_3, c_4, c_5, \dots$  be an enumeration of all contexts of usage  $c_i$  about which I can be asked: Does ‘red’ satisfy  $c_i$ ? If  $c_1 =$  ‘Fire-engines are red’,  $c_2 =$  ‘Tulips are red’,  $c_3 =$  ‘Lemons are red’,  $c_4 =$  ‘The planet Neptune is red’ then the function  $\mathbf{D}_{(CG,t,\text{red})}$  corresponding to my (CG) use of ‘red’ at time  $t =$  2pm, July 15, 2005 contains the following ordered pairs:  $\{\langle c_1, 1 \rangle, \langle c_2, 1 \rangle, \langle c_3, 0 \rangle, \langle c_4, ? \rangle\}$ .

---

<sup>96</sup> I borrow this rule from Crispin Wright [2007 p. 11].

(The use-contexts are not only sentences, as in this example, but can take a variety of forms, including, e.g., pointing to a red object and saying the word ‘red’.)

[J<sub>I</sub>] requires that  $\mathbf{D}_{(CG,t,w)}$  behave like a rule for any  $w$ , so that I can deduce a priori from  $\mathbf{D}_{(CG,t,w)}$  that I ought to apply  $w$  thus-and-so in any given use-context. It would behave like a rule-in-extension, at best (taking the form of a set of ordered pairs), if for any use context  $c_i$ , I can locate an ordered pair in  $\mathbf{D}_{(CG,t,w)}$  which contains  $c_i$  as argument and a determinate (yes/no) output as value. But since  $\mathbf{D}_{(CG,t,w)}$  is **partial** for most  $w$ , it will not always issue a determinate output: we’re often at a loss over whether a word satisfies a given use-context, due to, say, lack of information regarding the word, or some feature of the use-context. So clause (2) of [J<sub>I</sub>] is not satisfied by  $\mathbf{D}_{(S,t,w)}$  for most  $w$ . So, since for most  $w$ ,  $\mathbf{D}_{(S,t,w)}$  does not satisfy [J<sub>I</sub>], it cannot, according to the skeptic, serve as the sought after meaning-constituting fact. So premise (ii) of the justification argument appears to be correct.

### *Evaluating Premise (i)*

But should any candidate meaning-constituting fact be required to satisfy [J<sub>I</sub>]? Arguably, no. It is not difficult to see why. To begin with, *it’s not always the case that when I apply a term, I do so because I have some **reason** for thinking that it is the right application* – an explicit rule that I can produce on demand, or if someone were, say, to throw down a skeptical challenge. I often lack grounds for my knowledge in any useful general sense of the term, and reasons for following rules: I just follow them,

instinctively. So why should we force any candidate meaning-constituting fact – and in particular dispositional facts -- to satisfy [J<sub>1</sub>]? Moreover, why should we assume, as the *guidance* requirement in [J<sub>1</sub>+] demands, that I must consult any candidate meaning-constituting for guidance with each fresh application? This requirement *misconceives the phenomenology of our use of language*. How we are disposed to apply most words of our language is habitual, immediate, unreflective, and automatic. We don't *deliberate* over a particular course of action; it is triggered 'automatically', without the intervention of thought; 'we almost observe it happening',<sup>97</sup> 'as a matter of course' [PI §238]. Use of the notion of guidance involves an overly 'intellectual' conception of the phenomena of understanding – 'it underestimates the primitiveness of what is going on', as Colin McGinn puts it [1984 p. 16].<sup>98</sup>

We describe competence in the use of an expression as a matter of 'grasping a rule' or 'following a rule' that 'tells us what to do in each new case'. We think that our understanding of a word determines how it applies to an indefinitely large range of cases, even though at any given time we have considered only a small fraction of these. When we consider a new instance of application, we confidently judge that 'the rule we have all along had in mind' dictates whether we ought to apply the word or not. But we must

---

<sup>97</sup> Falk's expression [1994 p. 385].

<sup>98</sup> Ruth Millikan [1990] draws an interesting analogy with other biological organisms in defense of this point. Millikan likens 'following a rule' to a 'biological purpose'. Biological purposes, she says, are 'unexpressed purposes'; they involve 'purposefully conforming to an implicit or unexpressed rule'. [1990 p. 329] She contrasts this with 'expressed purposes', which involve 'purposefully following an explicit or expressed rule'. 'Biological purposes' are 'functions fulfilled in accordance with evolutionary design' [1990 p. 330], though they need not be innate. To illustrate, Millikan offers the example of the rule a hoverfly follows to catch its mate. It seems unlikely, she says, that the hoverfly calculates over any 'inner representation' of this rule in order to follow it. Rather, it has an 'unexpressed biological purpose to conform to the rule' [1990 p. 331] which consists in a 'genetically determined mechanism of a kind that historically proliferated in part because it was responsible for ... getting male and female hoverflies together.' [1990 p. 332]

beware of *inflating* these seemingly harmless ways of speaking, as the skeptic does, thereby creating a mythological picture of what it is to ‘follow a rule’. The skeptic assumes that understanding the meaning of a word consists in having explicit (identifying) knowledge of the rule I follow, an informative description that determines how I ought to apply the word in each fresh instance of application. This is where the chimera is created. This is where the skeptic’s identification with Wittgenstein’s interlocutor – the fellow who is perpetually tripped up by language – is most apparent. For there is no miracle-rule that does this work. In the grip of a false picture of what linguistic competence consists in, the skeptic reasons that any purported meaning-constituting fact must ‘tell me’ how I should apply a word in each new case, otherwise my applications are ‘arbitrary stabs in the dark’. And on this basis, he rules out dispositional facts as being meaning-constituting facts, since, as we have seen, facts about how I am disposed to apply a word do not ‘tell me’ how to apply a word in each new case, in the way the skeptic demands.

In some cases, I might be able to produce a rule of the kind the skeptic demands. For example, if asked to justify my responding ‘125’ when asked the sum of 68 and 57, I might say that I added 8 and 7 to get 15, that I put down 5 and carried 1, and so on. But then what if I am asked what my understanding of the more basic term involved in this rule -- ‘carrying’ -- consists in? etc. Ultimately, I am bound to reach a level where I cannot justify my application of a term by producing some informative rule that ‘tells me’ how to apply the term, or, still less, *recognize* some informative expression as the rule I follow. I will reach a level of **basic terms** where I act without any reason in terms of which I can justify my action, where I act unhesitatingly but ‘blindly’; a level of ‘brute

understanding’, as we might put it. In these cases, the agent’s behavior does not ‘depend on his having a grasp of some principle that goes beneath and could validate dispositions to particular responses.’<sup>99</sup> ‘Nothing takes place which can be regarded as *working out* what a rule requires ... one is *not aware of* any mediating process – of any route to the judgment which one might recover and cite by way of justification for it, -- but is just *smitten*, as it were, by the judgment.’<sup>100</sup> There is no informative rule here that figures as a premise in some process of reasoning I have gone through.

Terms for directly sensed properties, such as ‘red’, ‘blue’, ‘hot’, ‘cold’, ‘sweet’, ‘bitter’, ‘harmonious’, ‘dissonant’, or evaluative terms like ‘beautiful’, ‘good’, ‘just’, ‘right’ aptly illustrate this point. I might not be able to produce an informative rule that ‘tells me’ how to apply these terms. In the latter case, a certain object or event might create a certain feeling or mood in me (a confident inclination that this way is the *right* way to apply the term rather than another way), an unhesitating disposition to proceed in a particular way, which I cannot capture in an informative descriptive phrase. In the former case, were I pressed for a reason, I might say it looks like *that*, and *point* to the color of a red object (ditto for other terms from that category -- feels like *that*, sounds like *that*, tastes like *that*, smells like *that*). We will look in vain for any more substantial ‘explanation’. In Bertrand Russell’s jargon, the referents of these terms (‘red’, ‘hot’, etc.) are known only by *acquaintance* and not by *description*. It is not crucial to my mastery of them that there is something I could produce, or even recognize, as an informative

---

<sup>99</sup> Falk’s expression [1994 p. 382]

<sup>100</sup> Wright’s expression [2007, p. 10]

expression of the rule I follow.<sup>101</sup> My understanding here ('grasp of a rule') is best construed as consisting in **knowledge how** to use them as opposed to **knowledge that** they mean such-and-such (*propositional knowledge*), or **explicit knowledge** of how to apply them – knowledge that can be elicited with suitable enquiry or prompting (we leave it open how much prompting is allowable).<sup>102</sup> This is, as Jackman [2003 p. 34] or Barbiero [\*\*\*\*] point out, in keeping with current studies of the psychology of classification that suggest that we conceptualize experience in terms of *prototypes* rather than in terms of categories determined by sets of necessary and sufficient conditions. As Wittgenstein himself puts this point in PI §201, 'There is a way of grasping a rule which

---

<sup>101</sup> As Wright puts it, basic cases, where rule following is 'blind', are 'cases where rule-following is uninformed by anterior reason-giving judgment ... In such a case one follows a rule "without reasons" in the precise sense that one's judgments about the input condition for correct application of the rule are not informed by the exercise of concepts other than that which the rule concerns – that is, the concept whose expression the rule regulates and grasp of which consists in competence with that very expression.' [2007 pp. 16-17] 'Basic rule following, like all rule-following, is rational in the sense that it involves intentionality and a willingness to accept correction in the light of error. But that is not to say that it involves responsiveness to the requirements of the rule, conceived as instructions, as it were, which can feature in thought and rationally inform one's response. ... To say that in basic cases we follow rules blindly or without reasons is to say that our moves are uninformed by – are not the rational output of – any appreciation of *facts about what the rules require*. This is, emphatically, not the claim that it is inappropriate ever to describe someone as, say, knowing the rule(s) for the use of 'red', or as knowing what such a rule requires. Rather, it is a caution about how to understand such descriptions – or better: about how *not* to understand them. ... we should not think of the requirements of the rule as a state which *rationaly underlies* and enables competence, as knowledge of the rule for castling rationally underlies a chess player's successfully restricting the cases where she attempts to castle to situations where it is legal to do so. In basic cases there is no such underlying, rationalizing knowledge enabling the competence.' [2007 p. 18]

<sup>102</sup> I follow Michael Dummett [1991, p. 96] in characterizing 'explicit knowledge' in this way. It seems it would not even be appropriate to characterize the kind of knowledge I have about how to apply the given terms as **implicit knowledge**, which Michael Dummett defines as 'knowledge which its possessor is incapable, unaided, of formulating verbally, but of which he can recognize a formulation when presented with one.' [1991, p. 95] A subject has Implicit knowledge of a rule he follows provided that 'when he understands the statement of it, he acknowledges it as accurately describing his existing practice.' [1991, p. 96] Some of our knowledge of orthography and syntax, he notes, is of this kind. In these cases, 'there a possibility of eliciting assent to an explicit formulation, for bringing the speaker to recognize, not only that that which he is credited with knowing is true, but that it represents a principle that had been guiding his use of the language.' [1991, p. 96] To illustrate, he uses the nice example of someone who spells correctly but is unable to say when the final consonant of a verb is doubled before the termination '-ing': "why, for example, we write 'fitting' but 'crediting', 'referring' but 'proffering', 'summing' but 'consuming'; but when told that the consonant is doubled only when the final syllable is stressed and the vowel is short, he is likely, after a little reflection, to acknowledge that that is the rule he follows". [1991, p. 96]

is not an interpretation, but which is exhibited in what we call “obeying a rule” and “going against it” in actual cases’. Wittgenstein’s well known example of a *game* illustrates this point: the ways in which we employ this term may have no more unity than a rope has a single fiber, so that all I could do is point to *examples* of games – board games, card games, ball-games, Olympic games, etc. -- as opposed to a set of necessary and sufficient conditions (some exhaustive set of properties) that characterizes what all games supposedly share in common. ‘Don’t say: there *must* be something common, or they would not be called “games”’, Wittgenstein says in PI §66, ‘but *look and see* whether there is anything common to all. For if you look at them you will not see something that is common to *all*, but similarities, relationships, and a whole series of them at that...’ Not all games are amusing, nor is there always winning or losing or competition. Not all games require more than one player. Different parts are played by skill and luck. What we find is a ‘complicated network of similarities overlapping and crisscrossing’, such that we can say games form a family with ‘**family resemblances**’ [PI §67] – threads of common features running among the many things we call ‘games’ -- but have no single common trait. Asked to justify my use of the term ‘game’, all I may be able to do is point to examples and say ‘This and *similar* things are called games’. Understanding here is again best conceived as knowing *how* to use the term, which is exhibited in appropriate judgments about particular cases, as opposed to apprehending some manifest or latent essence (what all games supposedly have in common), as demanded by [J].<sup>103</sup>

---

<sup>103</sup> For further examples, consider expressions like ‘public interest’, ‘obscenity’, ‘intelligence’, ‘disease’, and ‘public disturbance’. No definition covers all the possible circumstances that such terms might cover. For example, as Norman Swartz points out for the term ‘public disturbance’: ‘The drafters of a law cannot be expected to foresee all the ways it is possible to disturb the peace. Courts of law constantly refine the

But if we reject [J<sub>I</sub>], as these observations urge us to, must we conclude that our applications of words are ‘arbitrary stabs in the dark’? No, though the skeptic would have us think otherwise. The skeptic attempts to force [J<sub>I</sub>] on the dispositionalist by presenting him with a **false dichotomy**:

Any purported meaning-constituting fact must justify our application of terms in the sense of [J<sub>I</sub>].

OR

Our applications are ‘brute inclinations’, ‘stabs in the dark’, ‘jack-in-the-box unjustified and arbitrary responses’. [pp. 10, 11, 15, 17, 18, 23, 55]

Of course our applications of terms are not just arbitrary ‘stabs in the dark’, except perhaps in the earliest stages of language learning. The stages of language learning for human beings with normal functioning cognitive capacities are, to put it very generally, from ‘no rule’ [NR] → ‘bent rule’ [BR] → ‘right rule’ [RR]. In acquiring a language, we

---

meaning of a term as new cases occur; they reduce the vagueness – and their decisions become precedents for future reference.’ <<http://www.sfu.ca/philosophy/swartz/conditions1.htm>>. A very nice passage illustrating how these decisions are made is quoted by Swartz from *Time* magazine (Canadian ed. [Jan. 1, 73], p. 34).

The state of Iowa plans to take legal action early next year against any citizen who would ‘allow, cause, or permit the emission of objectionable odors into the atmosphere.’ But what is the legal definition of an objectionable odor? The six members of the state’s air-pollution-control commission have been pondering that question, and at one meeting they even spent the better part of an hour debating whether an objectionable odor exists if no one is there to smell it. They decided that it does.

In more practical terms, the commissioners decided that an objectionable odor is legally objectionable if it lasts more than three hours, if it occurs more than once in three months, and if it is offensive to 30% of at least 30 randomly selected residents of an area.

Are these residents accurate judges? The commissioners tested themselves, sniffing wintergreen and amyl acetate (which smells like bananas) to see if they could tell the difference. They could. So they decided to give similar tests and to create a cadre of ‘trained noses.’ When a citizen complains about an odor, a jury of tested state employees will rush out to sniff and then decide if the law has been violated.

sometimes catch on to a bent rule. For example, in the early stages of language acquisition, children characteristically overgeneralize: they say *sleeped* instead of *slept*, *brang* (on the analogy of *sang*) instead of *brought*, and so forth. In these cases, they *are* following a rule, only it is the *wrong* rule. However, we hardly ever behave in accord with NR, except in the earliest stages of acquiring a first language (not as adults trying to master a second language). Infants, for example, tend to mouth words off arbitrarily; in such cases, they cannot be said to be following any rule. After a certain stage, NR no longer remains a plausible interpretation of a (normal cognitive functioning) person's linguistic behavior. One is forced down the route indicated by the arrows in virtue of being a member of a linguistic community. Mutual pressures towards conformity – the *training* process -- help adjust one's dispositions to correct patterns of usage.<sup>104</sup> For example, a young child might apply the term 'dog' to anything that has four legs. But if he calls a cat 'dog', he is corrected by others. Thus he comes to understand that although cats have four legs, they are not dogs. So he will revise his beliefs about what sorts of things are dogs, and eventually, after a series of revisions, will come to use the term 'dog' as experienced speakers of the language do.

The skeptic reasons that since the second horn of the dilemma is not a viable option, the first must be correct. But this conclusion rests on a *false dichotomy*, for the options do not exhaust the possibilities: there are ways other than [J<sub>1</sub>] in which our applications of words may be said to be justified. For example, an agent's application of word *w* may be said to be justified (in an *externalist* sense) if he applies *w* in ways that agree with how his linguistic community is disposed to apply *w*. The justification-conditions here need

---

<sup>104</sup> For a discussion of how conformity is generated through training, see Barbiero [\*\*\*\*].

not be explicitly known by the agent.<sup>105</sup> On this interpretation, which rule an agent is following is determined ‘from the outside’, not ‘from the inside’ (as the skeptic demands), by comparing the agent’s use of  $w$  to his linguistic community’s use of  $w$ , as follows:

- i.  $S$  is disposed to apply  $w$  in way  $\mathbf{D}_{(S,t,w)}$  in CP conditions.
- ii. The ‘community’ (standard-of-correctness setters) is disposed to apply  $w$  in way  $\mathbf{D}_{(\text{comm},t,w)}$  in CP conditions, and means  $M$  by  $w$ .
- iii.  $\mathbf{D}_{(S,t,w)}$  is similar to  $\mathbf{D}_{(\text{comm},t,w)}$ . (inductive inference).
- iv. Therefore,  $S$  means  $M$  by  $w$ .

A dispositionalist is entitled to assert (iv) only after having ascertained (i-iii). To say ‘ $S$  means  $M$  by  $w$ ’ is to make an inductive claim about the *degree of similarity* between  $S$ ’s dispositions in applying  $w$  and those of his linguistic community.

The skeptic’s attack on the dispositional theory misfires because of the unreasonable demand ([J<sub>I</sub>]) he makes on the **nature** and **role** of any purported meaning-constituting fact  $F$ , which translates into a set of *metaphysical* and *epistemological* constraints, respectively, on  $F$ .

---

<sup>105</sup> Alternatively, one might say that an agent is justified in applying  $w$  thus-and-so if he is disposed to *recognize* some non-trivial expression of a rule as correctly stating what he means. In particular, one might say that one is justified in answering ‘125’ to the query ‘68 + 57’ if one is disposed to recognize as correct some description of a rule for finding the sum of any two numbers. But this notion of justification can only apply to non-basic cases where a rule can be formulated.

**Metaphysical constraint:** *F contains instructions* that determine in advance how *w* is to be applied in an open-ended range of situations; thus *F* corresponds to a *complete* (fully defined) function.

**Epistemological constraint:** I have *direct access* to *F*, and I *consult* *F* with each fresh application of *w*.

But the function corresponding to how I am disposed to apply a word is, for *most* words of our language, one that is **progressively constructed/created** over time, and hence ‘gappy’. It is unreasonable to expect it to be *total*, still more, expressible as a rule which contains instructions for an open-ended range of situations, and, still more, as a rule I consult with each fresh application. Most words of our language, like the concept of a game, have open, not closed, boundaries (as we find in logic and mathematics). They are surrounded by ‘grey areas’ -- cases for which there are no rules of language that settle whether the term applies or not. In elucidating them, certain **models/paradigms** are usually given over which there can be no doubt or hesitation as to whether the word applies or not, but conditions of application cannot be *exhaustively* enumerated.<sup>106</sup> There

---

<sup>106</sup> The temptation to think otherwise arises, as Crispin Wright observes in his explication of Wittgenstein,

‘from the reflection that the explanation of a rule must eventually culminate in, or anyway ultimately be founded upon, the giving of illustrations of its application; and that any such illustrations are finite, and hence open in principle to an indefinite variety of interpretations. Yet explanations do usually, or so we suppose, secure mutual understanding. So somehow more is got across – the thought continues – than the pursuit of explanations can ever make completely explicit. Correct uptake of an explanation is having the right ‘something’ come into one’s mind as a result of the explanation; and the resulting informational state, though it is expressed in one’s subsequent practice with the concept concerned, essentially transcends it.’ [2001 p. 159]

The test for whether one has grasped the explanation depends on one’s ability to ‘go on’ to classify other things, not in the samples offered, in the ‘same way’. This may tempt one to describe the process of correct uptake as one that involves the formation of a rule that tells one how to go on in unexemplified cases.

will always be new situations that require a **decision/judgment** on our part to either extend the concept to cover the new case or exclude it from its extension. Where the new case is very similar to the paradigm, our response may be instinctive and automatic, ‘spontaneous in the way in which a good chess player may make a clever move without fully self-consciously rationalizing his grounds for it.’<sup>107</sup> But where there is a vague similarity between the new case and the paradigm, our response may be far from automatic and instinctive. To see this, just consider once again the case cited in the footnote above from *Time* magazine (Canadian ed. [Jan. 1, 73], p. 34): ‘The state of Iowa plans to take legal action early next year against any citizen who would “allow, cause, or permit the emission of objectionable odors into the atmosphere.” But what is the legal definition of an objectionable odor? The six members of the state’s air-pollution-control commission have been pondering that question, and at one meeting they even spent *the better part of an hour* [italics my emphasis] debating whether an objectionable odor exists if no one is there to smell it. They decided that it does.’ In forming such decisions,<sup>108</sup> we need to consider whether the new case is similar in enough important

---

Wittgenstein, as Wright notes, is concerned to expose this mistaken epistemological picture, and in turn ‘the rules as rails imagery which imply a figurative expression of Platonism’. [2001 p. 160] ‘The connection between the training and explanations which we received and our subsequent practices ... is not mediated by the internalization of explanation-transcendent rules that, in our training, we (something like) guessed’. [2001 p. 167] ‘It is, for epistemological purposes, a *basic* fact about us that ordinary forms of explanation and training do succeed in perpetuating practices of various kinds – that there is a shared uptake, a disposition to concur in novel judgments involving the concepts in question. The mythology of the “rules as rails” attempts an explanation of this fact. But the truth is the other way round: it is the basic disposition to agreement which sustains all rules and rule-governed institutions. The requirements which our rules impose on us would not be violated if there were not this basic agreement; they would not so much as *exist*.’ [2007 p. 7]

<sup>107</sup> Analogy used by Wright [2007 p. 17]

<sup>108</sup> Decision talk, unlike intuition talk (involved in apprehending latent essences), repudiates the suggestion that rule-following requires a ‘tracking’ (Crispin Wright’s term) epistemology; that we follow, as Wright puts it, ‘rules-as-rails, rules which somehow reach ahead of us and determine of themselves their every actual and counterfactual proper application’. [2001 p. 163] The rules-as-rails picture leads to insuperable difficulties, for: ‘We have no accountable idea of what would constitute the direction taken by the rule off

respects to the paradigm, so necessarily our response will not be automatic, but will require pause for thought. The point is, both  $\mathbf{D}_{(S,t,w)}$  and  $\mathbf{D}_{(comm,t,w)}$  will turn out to be partial functions for most  $w$ , progressively created as new cases arise and new decisions are made. It is in this respect that **meanings are constructed**, in piecemeal fashion, over time; they are hardly ever fully given. Construction of meaning is, to borrow a useful metaphor from McGinn [1984 p. 134] (though it is not one he endorses), like the existentialist construction of a self: it is the free action of an agent that constitutes what he is; it is not that action springs from an *antecedently* constituted self. Likewise, I submit, it is not that a rule is the *source* of a subject's use of a word (including how he is disposed to use the word), in the manner that [J<sub>I</sub>] requires; rather, exactly the opposite holds: how a subject is disposed to use a word determines the rule he follows (i.e. what he means by it), in the manner that (i-iv) above makes explicit.

*The fact that meanings are created has important implications for the 'underlying nature' of any candidate meaning-constituting fact.* As I like to view them, meanings are not stagnant, like dead matter, but alter and adapt to their environment with the passing of time, like living organisms. Those that fail to do so simply cease to exist. This analogy is pertinent, for, unlike complex forms of living matter:

- (a) Inorganic (dead) matter, like glass or a salt crystal, is disposed to behave in *determinate* ways in any given context. A salt crystal, e.g., does not decide over how to behave in any given environment. It does not hesitate over a certain course of action. Its behavior is, in some sense, *fixed in*

---

its own bat, as it were... once the direction is conceived, after the fashion of Platonism, as determined autonomously, and our performance, whether communal or solitary, as merely an exercise in tracking.

*advance* for any given context, though we might not know what it is. As science advances, so does this knowledge.

- (b) Its behavior in that context remains *fixed* over time.

By contrast:

- (a') How a subject is disposed to apply a word across a wide variety of contexts is not, for most words of our language, determined in advance (or at least, not with the same levels of accuracy), given that, as we have seen, many contexts require *decisions* whose outcome cannot be predicted in advance.
- (b') Moreover, unlike the behavior of inorganic matter, how a subject is disposed to apply a word across a wide variety of contexts is *may alter over time*.

So we must beware of *likening* dispositions associated with our use and understanding of language to dispositional properties of physical objects more generally, as the literature virtually always does!, for this identification blurs these important distinctions.

Dispositional properties of inorganic matter (like a salt crystal's disposition to dissolve in water, or the disposition of a piece of glass to shatter when dropped from a height) are *discovered*; they are *in them by nature*. This suggests that they are fully constructed, and not subject to change. By insisting that the fact F that constitutes meaning and understanding is some 'state' of an agent, and blurring the above

distinctions, the skeptic too assumes that F, whatever it may be, is *fully constructed*. But dispositions associated with how we are disposed to apply words are *created*; they *do not exist in us by nature*. I take dispositions associated with our applications of words to be **habits in the Aristotelian sense**. They are created in the same way that, for Aristotle, character traits are created: through regular *practice*. According to Aristotle [Book II, *Nicomachean Ethics*] we are not born courageous, temperate, just, honest, etc. Moral virtues are not implanted in us by nature. We acquire the moral virtues through practice. It is, e.g., by practicing the virtues of courage, honesty, justice and temperance, that we become courageous, honest, just, and temperate.

None of the moral virtues is implanted in us by nature, Aristotle says, because

- (i) If they were, then, of course, there would be no need to teach one how to be virtuous, and everyone would become virtuous by nature.
- (ii) Nothing that exists *by nature* can be changed by habit, so we would be stuck with our character traits. To illustrate, he says that it is impossible for a stone, which has a natural downward movement, to become habituated to moving upward, even if one should throw it up in the air ten thousand times to inculcate the habit.
- (iii) Of all the qualities we are endowed with *by nature*, Aristotle says that we are provided with the capacity first, and display the activity afterward. For example, it is not by frequent seeing or hearing that we acquire the ability to see or hear, but on the contrary, we first possess the ability and then use it; we do not acquire it by use.

Although the moral virtues are not implanted in us by nature, we are, Aristotle says, implanted with the *ability to acquire them*. And we do so through practice, as with other arts (men become builders by building houses, harpists by playing the harp, etc). Regular practice induces habits or dispositions. These bring our native ability to acquire moral virtues to completion and fulfillment. And habits or dispositions become our character traits.

In like manner, I submit, how we are *disposed to apply words* is something that is created through practice; it is not something that exists in us by nature. And thus, insofar as understanding a word consists in a disposition to apply it correctly, it too is something that is created through practice. If it were not this way, then of course we wouldn't need language teachers to train us to apply words correctly. Moreover, since nothing that exists *by nature* can be changed by habit, how we are disposed to apply words could not change over time. But this, of course, is not how it is. So, though in the sciences, dispositions are described as underlying physical 'states' of objects, which objects possess *by nature*, we must beware of regarding dispositions associated with applications of words as underlying physical 'states' in the *same* sense! Dispositions associated with applications of words are created through an ongoing process; they are not hardwired or permanently present.<sup>109</sup> And they are not merely biological, but shaped by external factors (a 'form of life'). They may alter over time, as I acquire new beliefs, or more accurate knowledge of words. Using my earlier notation, for most  $w$ , values in the range

---

<sup>109</sup> We are tempted to think otherwise because, as Wittgenstein points out 'we are strongly inclined to use the metaphor of something being in a peculiar state for saying that something can behave in a certain way. And this ... metaphor is embodied in the expressions 'He is capable of ...', 'He is able to multiply large

of  $\mathbf{D}_{(CG,t,w)}$  can alternate between ‘0’, ‘1’, and ‘?’ over time. By contrast, a stone’s disposition to fall to the earth when released from a height can never alter, unless the laws of gravity themselves alter. Likewise for a given species behavioral patterns: only a change in the environment is likely to cause a change in its behavioral patterns. So we must beware of *overstretching the analogy* with dispositional properties of physical objects more generally, since this includes objects whose behavior is far more stable and predictable than is our use of language.

This is not to say, for the interpretation I am offering of the *nature* of the dispositional facts corresponding to a subject’s use of words, that space cannot be made for the commonsense idea that words can be used wrongly, that what a subject now means by a word normatively constrains how he subsequently uses the word, and that if his subsequent use of a word is to be correct, then it needs to accord with or be faithful to what he now means by it. For certain contexts  $c_i$ , certain outputs are bound to *stabilize* at some point in the language acquisition process and remain stable over time. For example (to return to our earlier example), if  $c_1 =$  ‘Fire-engines are red’,  $c_2 =$  ‘Tulips are red’,  $c_3 =$  ‘Lemons are red’, then the function  $\mathbf{D}_{(CG,t,red)}$  corresponding to my use of ‘red’ at *any* time in the future is likely to contain the ordered pairs:  $\{ \langle c_1, 1 \rangle, \langle c_2, 1 \rangle, \langle c_3, 0 \rangle \}$ . For this subset  $\{c_1, c_2, c_3\} \subseteq \text{dom } \mathbf{D}_{(CG,t,red)}$ , and more generally for some subset of  $\text{dom } \mathbf{D}_{(CG,t,w)}$  for almost every word  $w$ , we can predict my behavior involving  $w$  in that set of contexts with high degrees of accuracy. In cases where there are ‘routines that are at work in us’,<sup>110</sup> we can say that how I am disposed to use  $w$  is in some sense ‘hardwired’.

---

numbers in his head’, ‘He can play chess’: in these sentences the verb is used in the present tense, suggesting that the phrases are descriptions of states which exist at the moment when we speak...’ [BB §66]

<sup>110</sup> Falk’s expression [1994 p. 393]

Repetition generates a substratum of **instinctive behavior** that doesn't involve reflection or the making of a decision/judgment. This point is also observed by Aristotle [Book II, *Nichomachean Ethics*] in the acquisition of moral virtues. Repeated acts of e.g. honesty, he says, make it easier to be honest, just as repetition in playing a musical instrument makes playing easier. Once one has acquired a virtue – for example, a disposition for honesty -- it becomes easier for that person to manifest the virtue – to be honest -- for it has become 'habitual' or 'second nature' to him or her.<sup>111</sup> Current neuroscientific studies show that *familiar use-contexts* are stored in the brain as synaptic weights that dispose us to *automatically* instantiate a given pattern of behavior. In these cases, we can say **causal mechanisms** are at work (though they are mechanisms that have been *created*) that trigger determinate courses of action. *When habits are formed, the phenomenology of my use of language changes, from deliberative and reflective to causal and instinctive.* The process is described by Barrie Falk [1994 p. 388-389] as follows: 'the current scene reminds me of an earlier one; I then recall how it developed and the moves I made in response to the development; and finally, I plan and launch into a sequence of moves designed to match the remembered response. No doubt, we sometimes do engage in such a procedure, but the effect of the past is normally simpler and more direct than this. Experience of regular processes in our environment and a history of (for the most part) successful attempts to manipulate it to our advantage affect present behavior by producing what are call habits.' Two crucial features of the operation of habits are (1) 'their presence can cause a current scene to prompt a sequential response to its

---

<sup>111</sup> The same applies, he says, to the opposite of virtue, namely vice. It is by practicing injustice, cowardice, or dishonesty, that we become unjust or cowards or dishonest. The person who lies and lies again finds that lying is easier and telling the truth more difficult. One can build *bad habits (vices)* as well as *good habits (virtues)*. Just like other bad habits, bad moral habits are difficult to change or break.

development, of a kind that normally leads to a routinely favorable outcome, without the agent's having to *infer* from his remembered response the set of moves that is now going to be required'<sup>112</sup>, and (2) 'I be prompted to the learned response by items that, as a matter of fact, are relevantly similar to the teaching paradigms. ... The agent does not think of the past but responds spontaneously to the current scene by launching into the routine that experience of scenes like it in the past has disposed him to.' So, though the function  $\mathbf{D}_{(s,t,w)}$  corresponding to a subject's use of  $w$ , for most  $w$ , is hardly ever complete and fixed over time, it enjoys enough stability<sup>113</sup> (through the creation of habits) to legislate most future applications of a word, so that he is not (normatively) free to apply words in an arbitrary fashion with complete linguistic impunity. The way I am suggesting we conceive of the *nature* of the dispositions associated with a subject's use of a word does not oppose the idea of future determination, but rather the mythical conception of what that determination consists in, as encapsulated in [J<sub>i</sub>]; in particular, it opposes the idea that in some queer way *all* future applications are already present in the meaning-constituting fact.

## 2.5 Conclusion

---

<sup>112</sup> Otherwise, Falk notes, it will not be possible to explain the unhesitating nature of the response. Rather, 'what I remember may sometimes cause me to resume the "current" of the activity I was then engaged in. That is to say, I do not just recall that I made the move and that the context in which I did so was thus and so: rather I re-experience the surrounding thoughts, feelings, desires in their role *as prompting me* to a particular response.' [1994 p. 396]

<sup>113</sup> In cases where my behavior is spontaneous.

Though the skeptical argument is flawed, it highlights/draws our attention to an important problem. As Kripke says, ‘much of the value of the sceptical argument consists precisely in the fact that he has shown that an ordinary practice, if it is to be defended at all, cannot be defended in a certain way’. [p. 66] The way it cannot be defended is, of course, as the skeptic demands: [J<sub>I</sub>].

The challenge the skeptical problem leaves us with, once this unreasonable constraint has been removed, is: find a meaning-constituting fact that navigates between the Scylla of an *overly rationalistic* conception of meaning and understanding (whereby I consult rules that guide my applications of words, as [J<sub>I</sub>] insists) and the Charybdis of an *irrational* ‘anything goes’ conception of meaning and understanding. These are the horns of the dilemma presented above.

I have tried to show in this chapter that [D<sub>CP</sub>] steers between the horns of this dilemma, since, on the one hand, I am not, according to [D<sub>CP</sub>], guided by rules, in the overly rationalistic sense of [J<sub>I</sub>]. Dispositions do not need to be interpreted; they are simply expressed, and *expression* is not an ‘interpretation’: it is not a propositional articulation of a rule.<sup>114</sup> [D<sub>CP</sub>] provides an intelligible way of clarifying the sense in which we think that our understanding of a word determines how it applies to an indefinitely large range of cases, even though at any given time we have considered only a small fraction of these. It provides a non-mysterious fact about an agent that has the property of ‘reaching out to’ or ‘containing’ all the absent cases not presently ‘before his mind’.

On the other hand, according to [D<sub>CP</sub>], how I am disposed to apply a term is not just a ‘shot in the dark’, an ‘unjustified and arbitrary response’, except in the very early stages

of language acquisition, since the *training process* helps bring one's dispositions in applying terms (so long as the subject has normal functioning cognitive capacities) in line with communal patterns of usage. Communal practices include tendencies to correct, criticize, and adjust deviations. So in a community, a deviant who calls a horse a 'cow', or makes mistakes in his calculations, etc. is criticized. In being corrected by others, he acquires new beliefs, and modulo these beliefs, how he is disposed to apply a term comes further in line with communal practices. Mutual pressures toward conformity, which form a basic part of the training process, help adjust one's dispositions to correct patterns of usage. Individuals, in virtue of their membership in a linguistic community, internalize these pressures and join in requiring these patterns of behavior in themselves and others. In this way, how we are disposed to apply a term stabilizes over time, so that, after a certain stage in the process of acquiring a language, the hypothesis that one's applications of words are 'jack-in-the-box unjustified and arbitrary responses' loses its plausibility.

---

<sup>114</sup> As Yalowitz puts it in [2000 p. 266].

## 3

## Wittgenstein's Philosophical Method

### CONTENT

---

3.0	Introduction
3.1	Symptoms
3.1.1	Some Remarks on Style
3.2	Diagnosis
3.2.1	How Pseudo-Problems are Generated
3.2.2	The Craving for Generality
3.2.3	The Moral Dimension of Wittgenstein's Philosophical Method
3.3	Treatment
3.3.1	Describing <i>Actual</i> Uses of Language
3.3.2	Exploring <i>Imaginary</i> Uses of Language

---

*It is a capital mistake to theorize before one has data. Insensibly one begins to twist facts to suit theories instead of theories to suit facts. -- Sherlock Holmes*

### 3.0 INTRODUCTION

In order to properly explain §§137-242 of PI, on which Kripke focuses his attention, we must situate them within the methodological framework in §§107-136 that precedes them. The purpose of this chapter is bring this methodological framework to light, and thus to set the stage for chapter 4, where I defend the claim that the skeptical problem *does* emerge in §§137-242, and that Wittgenstein treats it as just another pseudo-

philosophical problem, not in need of solution, but of dissolution. In this chapter, I try to clarify Wittgenstein's means of identifying pseudo-philosophical problems, and his means of treating them, by focusing mainly on his so-called 'meta-philosophical' remarks in §§107-136 of PI. There is an important analogy Wittgenstein develops in those sections between philosophy and medicine, which sees the practice of philosophy as a form of therapy, a method of healing. Uncovering the reaches of this analogy in Wittgenstein's work is one of the most fruitful and illuminating ways of explaining the aim and spirit of his philosophical method, and in particular, his diagnosis and treatment of pseudo-philosophical problems. I try to do so in this chapter by addressing the following questions, and defending the stated answers:

(§3.1)     Symptoms: What, for Wittgenstein, are the symptoms that are in need of philosophical therapy?

Proposed answer: Conceptual tensions.

(§3.2)     Diagnosis: How, for Wittgenstein, are the symptoms generated?

Proposed answer: From the urge to generalize across different 'regions of language'.

(§3.3)     Treatment: What is the nature of Wittgenstein's therapy?

Proposed answer: Describing actual and exploring possible uses of language.

### 3.1 SYMPTOMS

Wittgenstein wrote: ‘Working in philosophy ... is really more a working on oneself. On one’s own interpretation. On one’s own way of seeing things.’ [CV p. 16] One gets a feel for the truth of this by reading his philosophical writings. Not only are we drawn into a world of bizarre philosophical deliberations, but also, as Wittgenstein often describes it, a world of ‘personal struggle’. Many commentators would agree. For example, Stanley Cavell [1989 p. 37] writes:

The *Investigations* exhibits, as purely as any work in philosophy I know, a philosophizing as a spiritual struggle, specifically a struggle within the contrary depths of oneself, which in the modern world will present themselves as touches of madness.

and Rudolf Carnap [1967]:

When [Wittgenstein] started to formulate his view on some specific philosophical problem, we often felt the internal struggle that occurred in him at that very moment, a struggle by which he tried to penetrate from darkness to light ...

In what sense, for Wittgenstein, is work in philosophy ‘work on oneself’, and in what sense do his philosophical works exhibit an ‘internal’ or ‘spiritual’ struggle? It is important to address these questions if we wish to get clear on what, for Wittgenstein, is in need of philosophical therapy.

In attempting to answer them, I have found it helpful to make use of a well known expression coined by Sartre. In ordinary discourse, we sometimes say of a person that he shows signs of ‘bad faith’. According to Sartre, ‘bad faith’ is ‘a lie to oneself’, which we must distinguish from ‘lying in general’. The latter implies that the liar is ‘actually in

complete possession of the truth which he is hiding', 'affirms' the truth within himself. [1956 p. 87] Lying to oneself ('bad faith'), on the other hand, though it shares with lying in general the idea that the liar is presenting as truth an untruth, differs from lying in general in that it is *from himself* that the liar is hiding the truth. The deceiver, in this case, *is* the deceived. Bad faith thus results in a paradoxical duality of consciousness, an inner turmoil: the liar knows the truth, yet refuses to accept it, to act on it as a truth.

Though Sartre supposes that 'lying to oneself' involves *knowing the truth*, I think there is another way one can 'lie to oneself', in certain circumstances, that does not involve knowing the truth, and that is: to be ignorant of the truth yet *fail to seek it*. To interpret or see things as they are bequeathed to us by tradition when we are puzzled by those interpretations, troubled by them, perhaps even tormented by them, without an effort to understand *why* they trouble or torment us, is, in some sense, I would add, also to be in a state of 'bad faith'. Both this and the state of 'bad faith' Sartre speaks of involve an inner turmoil or unrest, and a *failure to act* in a way that aims at alleviating or dissolving that unrest.

Of course, it is not a failure to deal with *any* kind of problem generating internal tensions that, I would maintain, generates a state of 'bad faith'. Some problems might not have a *rational* explanation (e.g., Why did he behave that way?), so that efforts to resolve them must necessarily fail. Explanations do come to an end somewhere, and not always where we want them to. Fortunately, though, not everything that perplexes us is this way. For the most part, where there's a *will*, there's a *way* -- a means of rescuing ourselves from deception, prejudice and superstition; of enlightening ourselves.

At least, this is how Wittgenstein felt about many of the problems of philosophy. He too, though he did not employ these terms, had what I think can aptly be described as a good faith/bad faith distinction with regard to how we deal with what he took to be pseudo-philosophical problems. Wittgenstein felt that many of the problems of philosophy arise from a natural impulse or tendency to misconstrue the way language works, and thus are only pseudo-problems, conceptual muddles, that would fall like a ‘house of cards’ [PI §118] once their real nature is disclosed. *Good faith*, we can say, involves looking carefully into the ‘workings of language’ [PI §109], to get to the *source* of the pseudo-problems; *bad faith* a failure to do so. For only by ‘clarifying the use of our language’ [PG p. 115], Wittgenstein felt, can the philosopher remove particular misunderstandings that generate the pseudo-problems that ‘hold him captive’ [PI §115], like the fly in the bottle [PI §309]. The fly, having landed in a trap, is in grave danger. So too, Wittgenstein felt, is the philosopher when he fails to rescue himself from the snares of language (his prison).

I do not think it would be an exaggeration to say that for Wittgenstein, pseudo-philosophical problems can arise wherever *ordinary language* (common parlance) exists: they hover about (though they may go unrecognized) in virtually all disciplines that *borrow* expressions from ordinary language. And in this sense, his method of doing philosophy, which he likened to a form of therapy, has a *wide* range of patients. The practitioner of any discipline whose discourse overlaps with ordinary language can fall prey to the snares of language. Philosophy (that is, the practice of philosophy as Wittgenstein envisioned it) only appeared, for Wittgenstein, to be of no use to the ‘non-philosopher’, because of the traditional default of practitioners of various disciplines to

deal with (by means of eradicating) certain *kinds* of problems – specifically, for Wittgenstein, problems springing from *overstretched analogies* in our language (more on the nature of these problems later). As he says of the mathematician:

If a philosopher draws the attention of a mathematician to a distinction, or to a misleading mode of expression, the mathematician always says ‘Sure, we know all that, it isn’t really very interesting.’ He doesn’t realize that when he is **troubled by philosophical questions** it is because of those very unclarities that he passed over earlier with a shrug of the shoulders. [MS 219, 10]

A mathematician is bound to be horrified by my mathematical comments, since he has always been trained to avoid indulging in thoughts and doubts of the kind I develop. He has learned to regard them as something contemptible and, to use an analogy from psychoanalysis, he has acquired a revulsion from them as infantile. That is to say, I trot out all the problems that a child learning arithmetic, etc., finds difficult, the problems that education represses without solving. I say to those **repressed doubts**: you are quite correct, go on asking, demand clarification! [PG pp. 381-382]<sup>115</sup>

It is this dismissive way of pushing off a problem about a misleading mode of expression, suppressing a conceptual tension, or failing to resolve it, that gives rise to the ‘deep disquietudes’ Wittgenstein speaks of:

The problems arising through a misinterpretation of our forms of language have the character of depth. They are **deep disquietudes**; ...[PI §112]

A simile that has been absorbed into the forms of our language produces a false appearance, and this disquiets us. ‘**But *this* isn’t how it is!**’ – we say. ‘**Yet *this* is how it has to be!**’ [PI §112]

---

<sup>115</sup> No doubt, Wittgenstein had certain mathematicians or mathematically inclined philosophers in mind in writing this – Cantor, Russell, Ramsey, and Godel among them, all of whom he regarded as failing to recognize, or ignoring, the linguistic confusions that formed the bedrock of their ‘theories’.

And it is these ‘deep disquietudes’ [§112]<sup>116</sup>, repressed doubts or puzzlements, that I am assimilating to a state of ‘bad faith’, for they too trigger a duality of consciousness (‘*This* isn’t how it is!... Yet *this* is how it has to be!’ [§112]), a kind of *schizophrenic* disorder that is in need of therapy, only here the source of the conflict is language, as opposed to, as in the existential tradition, a denial of responsibility for our actions, or, as in the psychoanalytic tradition, a repression of painful childhood emotions.

Sartre maintains that though we are free beings, we are also quite ‘unaware’ of our freedom. This obliviousness results not from ignorance or oversight, but from the fact that we try to *conceal* our freedom from ourselves. But these efforts at self-deception, Sartre contends, are bound to fail, because human beings can try to conceal their freedom only to the extent that they recognize it. The attempt thus succeeds only in producing a paradoxical duality of consciousness / state of ‘bad faith’, since consciousness thinks of itself as a ‘thing’ (an entity which is not responsible for its behavior), yet at the same time gives recognition to its freedom (and hence responsibility).

Philosophers have also, Wittgenstein would say, generated clever tactics for concealing the real nature of pseudo-philosophical problems from themselves, treating them instead as genuine problems of profound difficulty: ‘Numerous traditions have treated this as a “serious” problem’, ‘Many philosophical theories have grown around it’, ‘It has occupied the minds of some of our greatest thinkers’, etc.

But efforts to solve pseudo-problems also give rise to the ‘deep disquietudes’ Wittgenstein speaks of. Once we recognize this, we are in a position to identify the *symptoms* -- the nature of the *kinds* of conflict -- that Wittgenstein saw as in need of

---

<sup>116</sup> Also described by Wittgenstein as a ‘mental uneasiness’ or ‘intellectual discomfort’ in Moore [1955, p. 27].

philosophical therapy. Just as an individual might go to a psychoanalyst or psychologist to bring rest to conflicting emotions he might have toward an individual, so too, Wittgenstein felt – and, I would hasten to add, recognized this better than any other philosopher of his time -- that we (and the philosopher in particular) can have conflicting attitudes / feelings over our understanding and use of expressions of our language (*'This isn't how it is! .... Yet this is how it has to be!'*), and this too is *in need of therapy*, generating, as it does, *internal (conceptual) tensions* – the 'deep disquietudes' Wittgenstein speaks of. This is why work in philosophy involves 'work on the self': it involves repairing a *fractured* self.



Though these tensions might not be *as intense* as conflicting emotions toward human beings can be, they do exist, and in some cases they are *very* intense. (Consider, for example, tensions concerning the 'existence' and 'goodness' of God, which can be a perpetual source of internal unrest for some theists.) By likening his philosophical method to the treatment of a disease, Wittgenstein was noticing the need for a 'science' to

treat the *kind* of ailment ('illness') that afflicts the philosopher (and non-philosopher as well<sup>117</sup>): *conceptual neurosis*. When we are suffering from conceptual tensions we have a bit of hidden or repressed nonsense in our minds, and the only way to cure it is to bring it out into the open -- to make it explicit nonsense. Wittgenstein saw his philosophical method as a therapeutic instrument for bringing conceptual tensions to the surface and dissolving them, just as Freud saw his psychoanalytic method as a therapeutic instrument for bringing conflicting emotions toward an individual to the surface and dissolving or at least alleviating them. In this respect, Wittgenstein's method is like psychotherapy. Like Freud, Wittgenstein's goal was *psychological health*: uniting a divided self.<sup>118</sup>

*How* conceptual tensions arise, and correlatively how, for Wittgenstein, pseudo-philosophical problems are generated, is a question I will turn to in §3.2. Before turning to that task, however, I think this is a suitable place to air some remarks about the *style* of Wittgenstein's writing.

---

<sup>117</sup> It is possible not only for the philosopher, but for the mathematician, the journalist, the scientist, the psychologist, the practitioner of virtually any discipline, to generate conceptual tensions, like those generated by the philosopher, and land himself in a muddle.

<sup>118</sup> On this note, we can also forge a connection between Wittgenstein and Socrates. Both sought to restore psychological health (for Socrates, health of the 'soul') in their interlocutors. Both saw the sickness of their time as consisting in a lack of self-knowledge. According to Richard Gilmore [1999], for the Socrates of Plato's so-called 'early dialogues', the lack of self-knowledge of his interlocutors manifested itself as 'a failure to acknowledge a disparity between their ways of acting and their ways of speaking and thinking'. [p. 141] This discrepancy resulted primarily from 'the mythology of the Homeric conception of the virtues to which was appended a newly developing conception of cooperation-based virtues'. [p. 142] This gave rise to a confused moral climate. On the one hand, we have Thrasymachus in Book I of the *Republic* professing that 'might makes right', and on the other, the moral principles of the *Euthyphro*, which are more closely aligned with cooperation-based virtues. [p. 142] What Socrates was fighting, on Gilmore's reading, was 'the temptation of people to follow popular conception', the 'spell cast by the opinion of the majority'; more specifically, 'a [Homeric] conception of the virtues that was outmoded', that had become 'a mythology'. [p. 139] The use of examples in Socrates philosophical method was to remind people of the *actual* contexts in which they use their words. [p. 142]

### 3.1.1 Some Remarks on Style

Wittgenstein's intense preoccupation with language, his recognition of the importance of investigating the various ways in which we put linguistic expressions to use, expressions seemingly unimportant and trivial, is, I think, more than anything else, what distinguishes his approach to philosophy from that of his predecessors -- his predecessors did not grasp the *extent* to which ordinary language is bound up in philosophical problems -- which is why, more than any other philosopher, Wittgenstein is responsible for its celebrated 'linguistic turn'. Indeed, Wittgenstein inaugurated a way of doing philosophy that was so revolutionary in manner and motives, and that influenced to so great an extent the practice of philosophy that was to follow, that it can be regarded as a *renaissance* in the history of philosophy, a kind of philosophical reformation.<sup>119</sup> Who ever would have thought it important, e.g., in attempting to *clarify* the notion of a 'mathematical proof', to observe that:

We can't say "I will work out *that* it is so", we have to say "*whether* it is so", i.e., whether it is *so* or otherwise. [PG]

A proposition of mathematics that is proved by an induction is not a "proposition" in the same sense as the answer to a mathematical question. [PG]

["how many are  $25 \times 16$ " and "what is  $\int \sin^2 x dx$ "] are "problems" in different senses. [PG]

A "question" in mathematics is not the same as what we call by that name in everyday life. [PG]

---

<sup>119</sup> According to G. E. Moore [1954], Wittgenstein himself felt this way about his work: '[Wittgenstein] said he was doing a "new subject", and not merely a state in a "continuous development of human thought", comparable to that which occurred when Galileo and his contemporaries invented dynamics; that a new method had been discovered, as had happened when "chemistry was developed out of alchemy" and that it was now possible for the first time that there should be "skillful" philosophers, though of course there had in the past been "great" philosophers.'

*Trying* to produce a phenomenon is not the same as *searching* for it. [PG]

If anything, these observations (only a handful of hundreds of a similar nature that pervade Wittgenstein's works) are bound to strike most people as *outbursts of insanity* -- the 'touches of madness' Cavell speaks of in the passage I quoted above. And yet, Wittgenstein deemed it of enormous importance to make them. Indeed, meticulous and painstaking investigations into the 'workings of language' revealed, for Wittgenstein, the *extent* to which ordinary language is bound up in philosophical problems: only by bringing into view trivial and well known facts about our use of language, Wittgenstein felt, can we untie the knots in our understanding that give rise to pseudo-problems. Trivial and seemingly inconsequential observations about how language functions have their point and place, though it may not be immediately visible (just as, in psychological counseling, it is not immediately obvious how bringing out seemingly insignificant details of one's past plays a role in the therapeutic process). This is why Wittgenstein says:

Doing philosophy is like opening a safe with a combination lock: each little adjustment of the dials seems to achieve nothing; only when everything is in its place does the door open.

And Moore, recalling his words:

He [Wittgenstein] also said that he was not trying to teach us any new facts: that he would only tell us trivial things -- 'things which we all know already'; but that the difficult thing was to get a synopsis of these trivialities, and that our 'intellectual discomfort' can only be removed by a synopsis of *many* trivialities --

that ‘if we leave out any, we still have the feeling that something is wrong. ...’<sup>120</sup>

Of course, we should not expect all Wittgenstein’s remarks to have a point and place. As he himself admits:

Only every now and again does one of the sentences that I write here make a step forward. [CV p. 66]

It seems to me that I am still a long way from understanding these things, a long way from the point of knowing what I do and what I don’t need to discuss. I still keep getting entangled in details without knowing whether I ought to be talking about such things at all; and I have the impression that I may be inspecting a large area only eventually to exclude it from consideration. [CV pp. 64-65]

We must take seriously Wittgenstein’s description, in the preface to PI, of his investigations as ‘long and involved journeyings’; in many cases, he may have led himself astray, so that his journeying leads to a dead end. But often his journeys only *appear* to lead to a dead end, because it is *we* who have been led astray by language, thereby mistaking relevant observations for irrelevant ones. The seeming distance of Wittgenstein’s linguistic investigations from the problem at hand is directly proportional to the length of the journey we must take through language to free ourselves from a false interpretation. This is why Wittgenstein says:

Why is philosophy so complicated? It ought, after all, to be completely simple. -- philosophy unties the knots in our thinking, which we have tangled up in an absurd way; but in order to do that it must make movements which **are just as complicated as the knots**. Although the result of philosophy is simple its methods for arriving there cannot be so.

The complexity of philosophy is not in its matter, but in our tangled understanding. [PR I, #2]

---

<sup>120</sup> Moore [1955 p. 27]

Wittgenstein's obsessive attention to seemingly irrelevant aspects of how language functions makes his works stand apart not only methodologically but also stylistically from other philosophical works. One striking aspect of his style of writing is the feeling of *eeriness* or *spookiness* it is bound to arouse in anyone who first encounters it. (I myself have a vivid recollection of this feeling upon first browsing through RFM). Why this is so is an interesting question, to be addressed in §3.3.2. It is also a style of writing that requires a great deal of patience from the reader. As Wittgenstein himself insists:

Sometimes a sentence can be understood only if it is read at the *right tempo*. My sentences are all supposed to be read *slowly*. [CV p. 58]

I really want my copious punctuation marks to slow down the speed of my reading. Because I should like to be read slowly. (As I myself read.) [CV p. 69]

Sometimes this insistence is presented in a more humorous light:

This is how philosophers should salute each other: 'Take your time!' [CV p. 81]

In philosophy the winner of the race is the one who can run most slowly. Or: the one who gets there last. [CV p. 34]

Of course, reading Wittgenstein is no garden stroll; it is more like finding yourself, as an archaeologist, amidst the shards of an ancient city (or, perhaps more aptly, to quote again from Cavell's remark above, 'within the contrary depths of [your] self'), where you are left to piece the fragments that surround you together yourself, to generate a coherent picture (a unified self). In this respect, reading Wittgenstein can be like reading the fragmented works of the Pre-Socratics. And if your journey is like mine, a great deal will remain scattered about, with a question mark hanging over it, and no apparent fit in some

larger whole. For Wittgenstein hardly ever tells us what the point of a remark is, or how it relates to other remarks. Conventional methods of guiding the reader are aborted. We have, instead, the impression of someone on the prowl, who has lost his way. We want to stop him and ask: What are you looking for? But Wittgenstein hardly ever reveals where he is heading. So in addition to (and intensifying) the feeling of eeriness that is generated in his works is a feeling of *caginess*, which has often been associated with the ‘poetic’ quality of his works.

There are, I think, at least three interrelated reasons for the caginess. One is that Wittgenstein sought to transform the nature of the reader, to force him into a kind of conversion, which crucially involved *unlearning* certain bad habits, so as to *demystify* pseudo-philosophical problems. This is why he says: ‘Working in philosophy ... is really more a working on oneself. On one’s own interpretation. On one’s own way of seeing things.’ [CV p. 16] And why, in the Preface to the *Philosophical Investigations*, he says ‘I should not like my writing to spare other people the trouble of thinking, but to stimulate someone to thoughts of his own’. Wittgenstein demanded *active participation* from his reader: the reader must work certain things out for himself, he felt, if his work is to have a *therapeutic* effect on him.

The caginess is also due to the fact that the aim of Wittgenstein’s investigations was not to *solve* what he deemed to be pseudo-problems but to *dissolve* them, to show why they are illegitimate. Pseudo-problems, Wittgenstein felt, originated from an impoverished diet of examples regarding the way words/sentences are used.

A main cause of philosophical disease – a one-sided diet: one nourishes one’s thinking with only one kind of example. [PI §593]

So to *deconstruct* the pseudo-problems, Wittgenstein felt it necessary to present a rich battery of examples that remind the philosopher of various uses of words/sentences. These include inventing language-games, comparing them, using diagrams and drawings, appealing to metaphors and jokes, etc. This style of writing creates a feeling of caginess because arrays of examples ‘criss-crossing in every direction’ [PI, Preface] lack the conventional guides associated with crisp arguments and linear progressions, as we typically find in philosophical works.

Finally, the caginess is also due to the fact that Wittgenstein felt that drawing out morals for the reader was *not necessary* for achieving his desired effect: demystification through the deconstruction of pseudo-problems. Wittgenstein allows the truth of what makes his philosophical method effective to speak for itself. Just as a work of literature (a novel) can produce moral effects not by *preaching*, but simply by *being*, so too, Wittgenstein felt, that he could achieve his aim – disintegration of pseudo-problems – by presenting a rich battery of examples, vividly described, of language-in-action, without having to *explicitly* draw any morals for us.

If only you do not try to utter what is unutterable then nothing gets lost. But the unutterable will be – unutterably – contained in what has been uttered.<sup>121</sup>

The moral potency of literature attests to this. In order for a work of literature to have moral effects, it is not necessary that it present a theory of morality.<sup>122</sup> Nor is it necessary

---

<sup>121</sup> In a letter from Wittgenstein to Paul Engelmann dated 9/4/1917. See Engelmann [1967].

that it portray morally idealized subjects following rigid rules. Indeed, its moral potency is likely to be greater when it portrays *morally mixed* characters, much like the average viewer (as you find in ancient tragedies), and vividly described crises and struggles through which they pass, *leaving us to draw the moral for ourselves*. As Aristotle masterfully recognized in the *Poetics*, the *idealized* subject, free of common faults, loses his/her ability to engage our attention, and in turn our sympathy; the one who is like us, who we can relate to emotionally, and hence sympathize with, is the *best* agent for imparting a moral message. So too, I believe, Wittgenstein recognized that the roughness and inexactness of commonplace speech, with all its imperfections, set before us in numerous examples, and not a neat philosophical theory, still more one couched in an idealized language, could serve as the *best* agent for producing the therapeutic (and arguably *moral*; see next section) effects he was after. And, for those effects to be produced, he felt, no morals need be explicitly stated. This contributes to the oracular, almost poetic quality of his work, and of course, once again, to the feeling of caginess.

## 3.2 DIAGNOSIS

### 3.2.1 How Pseudo-Problems are Generated

Throughout his life, Wittgenstein was preoccupied with investigating the ‘limits of language’. For Wittgenstein,

---

<sup>122</sup> This point has been defended at length by Martha Nussbaum in *The Fragility of Goodness* (Cambridge University Press, 1986), and *Love’s Knowledge: Essays on Philosophy and Literature* (Oxford University

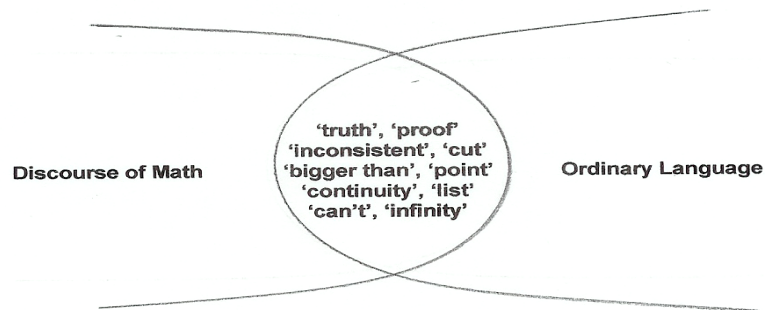
The results of philosophy are the uncovering of one or another piece of plain nonsense and of bumps that the understanding has got by running its head up against the limits of language. [PI §119]

A. Janik and S. Toulmin remark that the ancient dictum ‘Know yourself; Know your limits’ translated for Wittgenstein into ‘Know the limits of *language*’. [1973 p. 224] For Wittgenstein, one could obey the Socratic injunction ‘Know yourself’ only if he came to understand the scope and limits of his own understanding; and this meant, first and foremost, recognizing the precise scope and limits of language, which is the prime instrument of human understanding. [1973 p. 224]

But what is meant by the ‘limits of language’, a knowledge of which is integral to self-understanding, and, for Wittgenstein, to the practice of philosophy? By the ‘limits of language’, Wittgenstein meant, first and foremost, the limits of *our* language; that is, the language we actually use to communicate. Though our language is ‘in order’ as it is (i.e., connected to our ways of acting in the real contexts of our lives), it is, for Wittgenstein, the *source* of pseudo-philosophical problems. These problems arise in a number of ways, but I take the following to be central:

- (1) conflating *senses of words* in different ‘regions of language’, and
- (2) conflating *grammatical functions of sentences* in different ‘regions of language’.

(1) involves conflating the senses of words as they are used in different ‘practices’ or ‘language-games’. For example, conflating the sense of a word as it is used *within* the discourse/conceptual framework of science, math, religion, art, psychology, politics, etc. with how it is used *outside* the discourse; primarily, with how it is used in ordinary language. It’s quite natural to do this, since much of the terminology in these various fields is borrowed from ordinary language. For example, both within the discourse of mathematics and that of ordinary language we find the words:



Each of these words has a technical definition (specific to the discourse of mathematics), which (though resembling in some ways) differs in important respects from its ordinary language meaning. For example, in ordinary language, the word ‘infinite’ is rarely used to denote a quantity greater than every finite quantity (as it is *in* mathematics). Rather, the word ‘infinite’ is treated as if it were the designation of a *huge* number. We say, e.g., ‘I have an *infinite* amount of work to do!’ meaning a *huge* amount.

The fact that many expressions in a given discourse (‘language-game’) have a use both *within* and *outside* the discourse, in ordinary language, makes it tempting for someone who has not mastered the discourse, and, in some cases, even for someone who

has, to **conflate** the meaning of words within that discourse with their meaning in ordinary language. This can lead them to **falsely interpret** statements in the discourse embedding those words. False interpretations give rise to *conceptual tensions* ('This isn't how it is! ... Yet *this* is how it has to be!'): these are the *symptoms/signs* that something has been falsely interpreted. The unwary philosopher, seduced by the false interpretation ('This is how it has to be!'), is then led 'willy nilly' to erect what Wittgenstein took to be **pseudo-philosophical problems** (since they are based on false interpretations).

When we do philosophy we are like savages, primitive people, who hear the expressions of civilized men, put a **false interpretation** on them, and then **draw the queerest conclusions** from it. [PI §194]

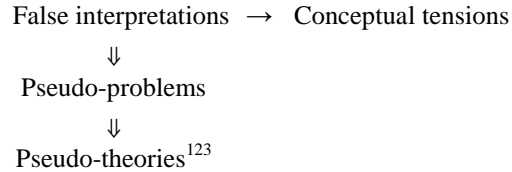
He then seeks to resolve these pseudo-problems in the wrong sorts of ways: constructing what Wittgenstein took to be **pseudo-theories** (thereby erecting a *new* home for the false interpretation).

(2) (which can give rise to (1) and vice versa) involves conflating *grammatical functions* of sentences in different 'practices' or 'language-games' -- functions as diverse as describing facts, commending, commanding, expressing feelings and emotions, influencing attitudes, etc. *Superficial similarities* in the syntactic form of sentences (e.g. the subject/predicate form) *conceal differences* in the role and function of those sentences. This can seduce the philosopher, once again, into raising pseudo-problems, which he seeks to resolve in the wrong sorts of ways (constructing pseudo-theories).

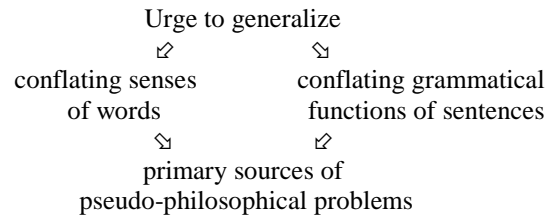
We can summarize the errors involved in (1) and (2) as follows:

Conflating senses of words or grammatical functions of sentences

↓



These tendencies to misconstrue how language works, which Wittgenstein saw as root causes of philosophical ‘sin’, can be traced in turn to the same fundamental urge: *the urge to generalize*. It is this urge that we can identify as the root source of the symptoms that Wittgenstein saw as in need of philosophical therapy. Pictorially, we have:



Wittgenstein emphasized that the puzzling questions that lead the philosopher to construct pseudo-theories are not in need of solution, but of **dissolution**: the philosopher needs to *draw his attention to the false interpretations* that lead him to posit the pseudo-questions in the first place, and recognize that the words embedded in the sentence(s) he falsely interpreted do not mean what he took them to mean, and/or that the sentences do not function the way he took them to function. To see this, the philosopher needs to examine how the (falsely interpreted) sentences function *within* the discourse/‘language-

---

<sup>123</sup> I follow Paul Horwich [2005] in referring to the problems generated by conflating senses of words as ‘pseudo-problems’, the mental tensions they generate as ‘conceptual tensions’, and efforts to solve them as ‘pseudo-theories’. Horwich offers a helpful discussion of how various fallacies – including overstretching analogies in our language – leads for Wittgenstein to the generation of pseudo-problems, in his [2005] chapter, which has influenced my own presentation above. His account of Wittgenstein’s metaphilosophy – merely sketched in his 2005 paper – was presented in much greater detail in various meetings of his Fall 2001 CUNY graduate seminar, which I attended.

game' that embeds them -- their natural surrounding / original home<sup>124</sup> -- *where they do their work*, not *outside* the discourse that embeds them, *where they remain idle* ('on holiday'). By doing so, the philosopher unties the knots in his understanding that give rise to the pseudo-problems, and the problems disappear.

Here are some examples to illustrate.

*Example 1*

Consider the meaning we attach to the notion of 'existence' in ordinary language. We think of an object of which we predicate existence (e.g. a chair or table) as *spacio-temporally bounded*. The ordinary use of this term can tempt one to *falsely interpret* statements *within* the discourse of mathematics involving the term, such as 'The set of natural numbers *exists*', 'The set of real numbers *exists*', as claims about *a completed totality, a finished product*. This interpretation of the infinite can then lead to *paradoxical results*, such as those found in set theory, which philosophers and logicians have attempted to resolve by constructing what Wittgenstein took to be *pseudo-theories* - - for example, *revising* the language of mathematics so as to remove all references to the infinite, and attempting to prove the consistency of the resulting theory. Wittgenstein noticed that there wasn't a need for such theories, for once it is recognized that the notion of a 'completed infinite totality' makes no sense, the paradoxes that prompted their construction would disappear. This is why he writes:

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

---

<sup>124</sup> The idea of bringing words home suggests mental economy, but also 'being at home with oneself', as Richard Gilmore [1999 p. 146] puts it; of restoring a self that has been fractured by language.

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). [PI §125]

A philosopher easily gets into the position of an incompetent manager who, instead of getting on with his own work and just keeping an eye on his employees to make sure they do theirs properly, takes over their work until one day he finds himself overloaded with other people's work, while his employees look on and criticize him. [CV p. 16]

*Example 2*<sup>125</sup>

We say, *within* the discourse of mathematics, that 'Numbers exist'. This sentence bears a resemblance to sentences in ordinary language, like 'Tables exist'. *Conflating* the functions of these sentences can lead the philosopher to *falsely interpret* the term 'number' to refer to an *object*, and the sentence 'Numbers exist' to function as a *descriptive statement*. This in turn can lead him to raise 'problems' that have an air of puzzlement -- what Wittgenstein took to be *pseudo-problems* (indeed, 'problems' that he thought lack sense<sup>126</sup>) -- like: What is the nature of these objects? Where are they located? How can we know anything about them? etc. He then attempts to resolve these 'problems' by constructing what Wittgenstein took to be *pseudo-theories*, since they address pseudo-problems.

One such mythological theory, for Wittgenstein, is *Platonism*. According to this theory, numerals are names of numbers, which are *abstract objects*. These objects are

---

<sup>125</sup> This example is due to Horwich [2005].

<sup>126</sup> This is why he says a philosophical problem has the form: 'I don't know my way about'. [PI §123] See also Moore 1955 p. 27: '[According to Wittgenstein, we are led] by instinct to ask certain questions, though we don't even understand what these questions mean.'

immaterial, not located, causally impotent. A ‘third world’, eternal, neither spatial nor temporal, is said to house them. Platonism accounts for our *knowledge* of abstract objects by positing a faculty of *intuition*, which puts us in contact with them. This faculty is supposed to be like sense perception, but also in some mysterious way different from it.

Wittgenstein emphasized that the puzzling questions that led the philosopher to create this pseudo-theory are not in need of solution, but of *dissolution*: the philosopher needs to *draw his attention to the false interpretation* that led him to posit the pseudo-questions in the first place, and recognize that numerals do not pick out objects in the world in the way that names of physical objects do. To recognize this, the philosopher needs to examine the *role and function* of the sentence ‘Numbers exist’ *within* the discourse of mathematics, *where it does its work*, not *outside* the discourse, *where it remains idle*. By doing so, the philosopher unties the knots in his understanding that gave rise to the pseudo-problems, and the problems *disappear*.

### 3.2.2 The Myth of Platonism

Insofar as philosophical ‘problems’ are based on false interpretations of the meaning and function of words and sentences of our language, Wittgenstein would dismiss them as pseudo-problems, and the theories they give rise to as pseudo-theories. It doesn’t follow that Wittgenstein thought that *all* philosophical problems are pseudo-problems, based on false interpretations of language, though many commentators have attributed this *radical* thesis to him. It would be quite strange if he held such a view, since there appear to be

innumerably many philosophical ‘problems’ that do not derive from false interpretations of the meaning and function of words/sentences of our language. Consider: What is art? justice? responsibility? friendship? happiness? etc. Arguably, these problems do not -- unless falsely interpreted as, e.g., scientific problems -- arise from conflating meanings of words or grammatical functions of sentences (unlike the questions: ‘Where are numbers located?’, ‘What are they made of?’); rather, one might argue, they arise from a *practical* need, e.g., the need to determine what kinds of objects to house in an art museum, how wealth ought to be distributed, what form of government to institute, when ‘criminals’ should be punished, how to pick your friends, how to live the best life, etc. Indeed, it was, in part, the practical question of how to create a just state that incited Plato to seek an answer to ‘What is justice?’ in the *Republic*, and the practical question of how to achieve ‘union and brotherhood’ among man that incited Tolstoy to seek an answer to ‘What is art?’ in *What is art?*, etc.

Wittgenstein would not dismiss the stated problems as pseudo-problems. Nor would he say (as he is commonly held to) that the best *we* can do by way of answering them is describe how the terms in question are used in *our* linguistic community (suppose, e.g., we live in an unjust state, or a state that glorifies trivial, banal, unwholesome ‘art’ – what Tolstoy called ‘counterfeit’ or ‘degenerate art’). Wittgenstein says in PI §182: ‘the role [of words] is what we need to understand in order to resolve philosophical *paradoxes*’ [italics my emphasis]; he doesn’t say: ‘the role [of words] is what we need to understand in order to resolve philosophical *problems*’. The point is that we need to turn our attention to how words are used in order to *deconstruct* pseudo-philosophical problems, not in order to *solve* philosophical problems. It is also a mistake to think Wittgenstein

denies (along with some radical relativists) that there can be better or worse answers to the stated questions. Some theories of, say, art, justice, the nature of God, etc., he would agree, can be said to be better than others, but *relative* to the work we want them to do for us -- the aim/goal we wish to achieve with them -- and this will be determined by the 'language game'/'form of life' which embeds them. The 'game' analogy is very telling. Games have *goals*. We can speak of whether a move is *warranted* in a game, of the *role* and *function* of making a move, and of whether it is a *good* move. We can *compare* moves, and speak of one move as being *better* than another, etc. In so far as we view a theory of x as a recommended definition for achieving the goal(s) of the game which embeds x, then it is possible to speak of better or worse theories of x by debating the reasons offered for the facility of one set of definitional criteria over another. Suppose, e.g., the goal of our game is to maximize happiness for all players. Then we can view competing theories of, say, 'justice', as competing criteria for achieving the goal of maximizing happiness. We can't speak of the 'best' theory of x, unless all games that embed x share the same goal, but we can speak of the 'best' or 'better'/'worse' theories relative to some game. Words, like tools, have a functional purpose. And they can function in different ways, depending on what we want to achieve with them, the work we want them to do for us. They don't carry the same function into every context. And their meaning is defined by their function. So naturally they will have different meanings in different contexts. In examining the question 'What is art?', Morris Weitz<sup>127</sup> makes an observation with which I think Wittgenstein would wholeheartedly agree: 'What is of utmost importance is the reasons proffered ... for the respective theory, that is, the

---

<sup>127</sup> Morris Weitz [2003], p. 783.

reasons given for the chosen and preferred criterion ... The value of each of the theories [of art] resides in its attempt to state and to justify certain criteria which are either neglected or distorted by previous theories. ... the role of a theory is not to define anything but to use the definitional form, almost epigrammatically, to pinpoint a crucial recommendation ...' If we construe the *motives* that lie behind the formulation of many traditional philosophical theories (proffered answers to 'What is x?' questions) as recommendations for adopting one set of criteria over another for achieving some goal, then we can see Wittgenstein as dealing more generously with them (instead of regarding them, as is commonly conceived, as worthless and doomed to failure). For incorporated in each of them, as Weitz observes, is a serious debate over and argument for emphasizing certain criteria over others.<sup>128</sup>

What Wittgenstein *would* deny is that the concepts mentioned above -- art, justice, responsibility, happiness, friendship -- and other fundamental concepts that have been explored by philosophers -- truth, freedom, meaning, causation, goodness, time, etc. -- have *hidden essences* (meanings that they purportedly carry with them into *every* context/game), a thought which is easy to lapse into if we treat 'What is x?' questions involving these concepts as scientific or metaphysical, or as assuming there is an exhaustive list of properties shared in common by all instances of 'x'. For Wittgenstein, the aforementioned questions are legitimate, so long as they are not construed as concerned with latent essences. Wittgenstein was opposed to treating philosophical problems as:

---

<sup>128</sup> Ibid, p. 783.

- (1) *analytic* ones, where an essence is taken to be a set of necessary and sufficient conditions governing a word's application.
- (2) *metaphysical* ones, where an essence is taken to be an abstract object, housed in some mysterious supra-sensible realm (e.g. *platonism* in the philosophy of mathematics, *psychologism* in the philosophy of mind).
- (3) *scientific* ones, where an essence is characterized in terms of the vocabulary of physics.

Though for scientific questions of the form: 'What is x?' (e.g. What is water?), offering a *physical theory* might be appropriate, it was not, for Wittgenstein, appropriate for *philosophical* questions, many of which are 'What is x?' questions (though the superficial similarity of these questions might tempt us to think otherwise). And so, Wittgenstein would dismiss as incoherent the imperialistic view of philosophy as the 'Queen of the Sciences', whose mission it is to uncover the fundamental nature of our most fundamental concepts.

Philosophers very often talk about investigating, analyzing, the meaning of words. But let's not forget that a word hasn't got a meaning given to it, as it were, by a power independent of us, so that there could be a kind of scientific investigation into what the word really means. A word has the meaning someone has given to it. [BB p. 28]

[The traditional questions] see in the essence, not something that already lies open to view and that becomes surveyable by a rearrangement, but something that lies beneath the surface. Something that lies within, which we see when we look into a thing, and which an analysis digs out.

'*The essence is hidden from us*': this is the form our problem now assumes. We ask: 'What is language?'. 'What is a proposition?' And the answer to these questions is to be given once and for all; and independently of any future experience. [PI §92]

The point is: *meanings are created*, in a variety of ways and in a variety of language-games, and they are created to satisfy a variety of needs, some practical, some theoretical.

There are no overarching meanings or essences that all words possess, common to all

language-games in which they are employed, awaiting discovery. Most of our words have *open* boundaries, in the sense that necessary and sufficient conditions for their application cannot be exhaustively specified, as Wittgenstein aptly illustrated with the term ‘game’. The ways in which we employ this term may have no more unity than a rope has a single fiber, so that all I could do is point to *examples* of games – board games, card games, ball-games, Olympic games, etc. -- as opposed to a set of necessary and sufficient conditions (some exhaustive set of properties) that characterizes what all games supposedly share in common. ‘Don’t say: there *must* be something common, or they would not be called “games”’, Wittgenstein says in PI §66, ‘but *look and see* whether there is anything common to all. For if you look at them you will not see something that is common to *all*, but similarities, relationships, and a whole series of them at that...’ Not all games are amusing, nor is there always winning or losing or competition. Not all games require more than one player. Different parts are played by skill and luck. What we find is a ‘complicated network of similarities overlapping and crisscrossing’, such that we can say games form a family with ‘**family resemblances**’ [§67] – threads of common features running among the many things we call ‘games’ -- but have no single common trait.

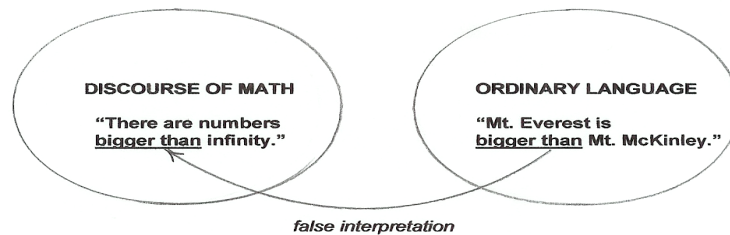
The philosophical thought that there are hidden essences awaiting discovery is just another facet of the *urge to generalize*. Here the urge surfaces in the assumption that there is something that different uses of words or sentences have *in common* (an essence).

We are inclined to think that there must be something in common to all games, say, and that this common property is the justification for applying the general term ‘game’ to the various games; whereas games form a *family* the members of which have family likenesses. Some of them have the same nose, others the same

eyebrows and others again the same way of walking; and these likenesses overlap. The idea of a general concept being a common property of its particular instances connects up with other primitive, too simple, ideas of the structure of language. It is comparable to the idea that *properties* are *ingredients* of the things which have the properties; e.g. that beauty is an ingredient of all beautiful things as alcohol is of beer and wine, and that we therefore could have pure beauty, unadulterated by anything that is beautiful. [BB p. 17]

The primitive forms of our language – noun, adjective, and verb – show the simple picture to which it tries to make everything conform. [MS 213, 425]

But it is precisely the urge to generalize across different ‘regions of language’ (as in the diagram below) that (as we saw in the last section) is the source of pseudo-philosophical problems.



What the philosopher has to do then, to avoid the conceptual muddles that give rise to pseudo-problems, for Wittgenstein, is learn how to temper the urge to generalize: to dissociate what should be kept apart.

### 3.2.3 The Moral Dimension of Wittgenstein’s Philosophical Method

To wrap up, the source of the symptoms that Wittgenstein felt were in need of philosophical therapy can be traced, at root, to an unbridled urge to generalize. Allowing this urge free rein leads to the kind of intellectual dishonesty involved in generating pseudo-philosophical problems and pseudo-theories. For it is mainly this that generates pseudo-philosophical problems: a failure to keep the various senses and functions of words/sentences apart in our minds. Mental sloth/laziness drives the philosopher to *assimilate* what should be kept apart ('as if the sense were an atmosphere accompanying the word, which it carried with it into every kind of application' [PI #117]). In our earlier examples, we saw how this causes the philosopher to generate pseudo-problems.

Of course, it is not only the philosopher who succumbs to such sloppiness: no one is completely immune from making such mistakes. Language, Wittgenstein felt, operates in 'dangerous' ways not only in the life of the philosopher, but in our daily lives, as the following remark made to Norman Malcolm [1962 p. 30] suggests (recorded by Malcolm in his biography of Wittgenstein):

Whenever I thought of you I couldn't help thinking of a particular incident which seemed to me very important ... you made a remark about the 'national character' that shocked me by its primitiveness. I then thought: what is the use of studying philosophy if all that it does for you is to enable you to talk with some plausibility about some abstruse questions of logic, etc., & if it does not improve your thinking about the important questions of everyday life, if it does not make you more conscientious than any ... journalist in the use of the DANGEROUS phrases such people use for their own ends.

Every one of us, whether we choose it or not, given certain features of our language, is liable to fall prey to the *kinds* of 'problems' that hold the unwary philosopher captive.

Language sets everyone the same traps; it is an immense network of easily accessible wrong turnings. And so we watch one man after another walking down the same paths and we know in advance where he will branch off, where walk straight on without noticing the side turnings, etc. etc. What I have to do then is erect signposts at all the junctions where there are wrong turnings so as to help people past the danger spots. [CV p. 18]

This is the *moral demand* that Wittgenstein makes, not merely, as the passage indicates, on himself, but, through his writing, on us: that we ‘erect signposts’ to help ourselves past the ‘danger spots’, that we remain in constant vigilance of the seductive and overwhelming powers of language, that we be *honest* with our selves when we are in the grip of a conceptual tension (‘*This* isn’t how it is ... But *this* is how it has to be!’), and look carefully into the ‘workings of language’ [PI §109], to get to the *source* of it. We must ‘distrust language’ [NB] in the sense of guarding from an instinctive or habitual urge to generalize, and reflect more carefully on the workings of language when we are confused.

Of course, this is not easy work:

‘Philosophy is a **battle** against the bewitchment of our intelligence by means of language’. [PI §103]

The philosopher is ‘engaged in a **struggle** with language’. [CV p. 12]

The struggle involves not so much a battle of the *intellect*, as of the *will*: the will to resist an unguarded urge to generalize (more specifically, ‘urge to misunderstand’ [§109]), one of the most deep-seated human urges.

What makes a subject hard to understand – if it’s something significant and important – is not that before you can understand it you need to be specially

trained in abstruse matters, but the contrast between understanding the subject and what most people *want* to see. Because of this the very things which are most obvious may become the hardest of all to understand. What has to be overcome is a difficulty having to do with the **will**, rather than with the intellect. [CV p. 17]

As I have often said, philosophy does not call on me for any sacrifice, because I am not denying myself the saying of anything but simply giving up certain combination of words as senseless. But in another sense philosophy demands a renunciation, but a **renunciation of feeling**, not of understanding. Perhaps that is what makes it so hard for so many people. It can be as hard to refrain from using an expression as it is to hold back tears or hold in anger. [MS 213, 406]

What we need to do, then, is strengthen the will to resist the urge to misunderstand. It is precisely here that we can extract moral content from the method Wittgenstein urges us to employ in our philosophical practice, or the practice of interpreting language more generally. For it is a failure to restrain an unguarded urge to generalize that is a root source, not only of philosophical error ('sin', to use Wittgenstein's expression), but of moral sin more generally. (Haven't, for example, hasty generalizations concerning what others are like been a root source of religious wars, racial prejudice, and other forms of oppression?) For Wittgenstein, God is in the details, and if we wish to save our selves from falling into certain kinds of error ('sin'), we must attend to the details.

Anthony Kenny [1982] compares Wittgenstein's conception of 'philosophical sin' to the Christian doctrine of original sin. 'Philosophical sin' is not something we are born with, Kenny says, it is something we take in along with language. 'Along with language, along with all the benefits which language brings, along with all the possibilities for our way of life which it brings, we take in whether we want to or not, certain temptations; we must resist these if we are not to be misled.' [1982 p. 15] This is why we are better off having gone through philosophy: we have gone through a discipline that helps us develop

*good habits*<sup>129</sup> to counteract the bad habits that we have acquired through our immersion in language.

Developing good habits, of course, is not easy work, for the grip of a false interpretation is very difficult to shake off once that interpretation has become *sedimented in our 'forms of life'*: to the extent that it has become sedimented, it becomes more difficult for us to challenge it.

Getting hold of the difficulty *deep down* is what is hard. Because if it is grasped near the surface it simply remains the difficulty it was. It has to be pulled out by the roots; and that involves our beginning to think about these things in a new way. The change is as decisive as, for example, that from the alchemical to the chemical way of thinking. The new way of thinking is what is hard to establish. [CV p. 48]

Wittgenstein would have agreed full heartedly with George Orwell when he wrote: 'An effect can become a cause, reinforcing the original cause and producing the same effect in an intensified form, and so on indefinitely. Language becomes ugly and inaccurate because our thoughts are foolish, but the slovenliness of our language makes it easier for us to have foolish thoughts.'<sup>130</sup> In this way, a mythology becomes 'embedded in our language' [MS 213, 434]. Moreover, we drill ourselves so often and so make a *habit* of interpreting words or sentence-forms in terms of standardly assigned

---

<sup>129</sup> Wittgenstein recognized that not everyone is in a position to develop/nurture such habits. As Anthony Kenny remarks '[For Wittgenstein] only those who were not at home in the world [language], those who found that they had to push against it, really had a hope of salvation.' [1982 p. 16]. Kenny cites the following passages in support of his claim:

Human beings are profoundly enmeshed in philosophical – i.e., grammatical – confusions. They cannot be freed without being extricated from the extraordinary variety of associations which hold them prisoner. You have as it were to reconstitute their entire language. – But this language grew up as it did because human beings had – and have – the tendency to think in this way. So you can only succeed in extricating people who live in an **instinctive rebellion** against language; you cannot help those whose entire instinct is to live in the herd which has created this language as its own proper mode of expression. [MS 213, 423]

meanings/functions that we *naturally* stretch these interpretations into all contexts where these expressions or sentential-forms are found, and until something compels us to reconsider our interpretation, we cease to think twice about it. In this state of contentment, we fail to notice, and so take steps to free ourselves from, the forces of language that hold us captive; our complacency conceals our shortness of vision.

A *picture* held us captive. And we could not get outside it, for it lay in our language and language seemed to repeat it to us inexorably. [PI §115]

But notice that Wittgenstein says we ‘*could not* get outside it’, not ‘*cannot* get outside it’. And indeed, the point (as the past tense ‘could not’ suggests) we must not forget is: the processes that have led to our bewitchment are *reversible*. Though philosophical writing has its share of bad habits that spread by imitation and tradition, these habits can be reversed, by *unraveling* the linguistic confusions that led to them.

... philosophy unties the knots in our thinking, which we have tangled up in an absurd way; but in order to do that it must make movements which are just as complicated as the knots. [PR Part 1 #2]

By tracing the usage of the concepts that we have ‘tangled up in an absurd way’, it becomes possible to see how they became entangled and thus to disentangle them.

The fact that bad habits can be *reversed* highlights the *moral dilemma* that confronts us: Do we follow long-standing conventions and traditions in our interpretations (‘*This is how it has to be!*’), when we are troubled or puzzled by those interpretations, or shall we courageously question the bulwark of popular opinion (‘*This can’t be how it is!*’), when

---

<sup>130</sup> *A Collection of Essays* (Harcourt Brace and Company, 1981) pp. 156-7

we are troubled by it, and probe into the ‘workings of language’ for possible sources of confusion. For Wittgenstein the answer is clear:

What I do think essential is carrying out the work of clarification with COURAGE: otherwise it becomes just a clever game. [CV p. 19]

In the next and final section, I will discuss the nature of Wittgenstein’s methods of unraveling the processes that have lead to philosophical bewitchment, and thereby, his means of treating pseudo-philosophical problems.

### 3.3 TREATMENT

1930 marked a turning point in Wittgenstein’s intellectual development. When he returned to Cambridge in that year, he thought, at last, that he discovered a *method* of doing philosophy that would allow him to treat all philosophical problems with professionalism and efficacy.

The nimbus of philosophy has been lost. For we now have a method of doing philosophy, and we can speak of *skillful* philosophers.<sup>131</sup>

Yes, I have reached a real resting place. I know that my method is right. My father was a businessman, and I am a businessman: I want my philosophy to be business-like, to get something done, to get something settled.<sup>132</sup>

---

<sup>131</sup> *Lectures* 1930-32, p. 21.

<sup>132</sup> Rhees [1984, p. 110].

There are two aspects to Wittgenstein's method of deconstructing pseudo-philosophical problems that need to be distinguished: (1) describing *actual* linguistic practice, and (2) constructing *hypothetical* 'language-games'. In a word, Wittgenstein's method consisted in the investigation of uses of language, *actual* and *possible*. Both methods were, for Wittgenstein, an indispensable means of clarifying the 'grammar' of our words -- i.e., the appropriate contexts for using those words -- and thereby dissolving pseudo-philosophical problems. Though (2) is often conflated with (1), it is important to recognize that it differs from it in important respects. (1) can be seen as functioning as a *direct* method of 'proof' (attempt to convince the reader of some thesis), and (2) as an *indirect* method of 'proof' -- *proof by reductio*. I try to clarify this distinction in what follows.

### 3.3.1 Describing *Actual* Uses of Language

(1) involves describing and comparing concepts as they are used in their *actual* linguistic settings, noting that 'Yes, this is what we *say*, and also *this*, and sometimes *that...*', reminding ourselves of uses of language that we have forgotten or overlooked. By reminding the philosopher of the role and function of expressions of our language, Wittgenstein hoped to weaken the hold that misleading analogies (examples of which I gave in §3.2.1) have on him. For many of the 'problems' that arise in philosophy, Wittgenstein recognized, rest on overstretched analogies, which the philosopher can dissolve by reminding him/her-self of how certain expressions are employed in their *actual* contexts of usage.

... the sentence only seems queer when one imagines a different language-game for it from the one in which we actually use it. [PI §195]

Describing the various ways we *actually* use words helps relieve us from a strained way of looking at them. This is the point of Wittgenstein's remark that:

A human being is imprisoned in a room, if the door is unlocked but opens inward; he, however, never gets the idea of pulling instead of pushing against it. [RFM III, 37]

This is a recurring image (imprisonment) Wittgenstein had of the philosopher, who he felt he must rescue. Pushing against the door is like the philosopher's treating philosophical problems like scientific problems: 'We still don't know that; but it is knowable and it is only a matter of time before we get to know it!' [CV p. 40] For Wittgenstein, this is the wrong attitude to take to philosophical problems:

It was true to say that our considerations could not be scientific ones. ... And we may not advance any kind of theory. There must not be anything hypothetical in our considerations. **We must do away with all explanation, and description alone must take its place. And this description gets its light, that is to say its purpose, from the philosophical problems.** These are, of course, not empirical problems; they are solved, rather, by looking into the workings of our language, and that in such a way as to make us recognize those workings: *in spite of* an urge to misunderstand them. The problems are solved, not by giving new information, but by arranging what we have always known. [PI §109]

**Philosophy simply puts everything before us and neither explains nor deduces anything.** – Since everything lies open to view there is nothing to explain. For what is hidden, for example, is of no interest to us. [PI §126]

It is important not to misunderstand what Wittgenstein is saying here, as commentators often have. There is a way in which I think he would grant that philosophy *is* a science, and *does* construct theories. One way to think of a scientific theory is as a statement that organizes a set of concepts in a meaningful way by explaining the relationship among them. The various sciences assume there is some underlying order in the universe; that events, whether they involve molecules or human beings, are not haphazard. They follow a pattern that is sufficiently regular for generalizations to be made about them. Generalizations are crucial to science because they place isolated, seemingly meaningless events in patterns we can understand. It then becomes possible to *explain* why something happens and to *predict* that it will happen again under the same conditions. In addition to the natural sciences, which try to explain *physical and biological phenomena* in this respect, and the social sciences (sociology, economics, psychology, political science, and anthropology) which try to explain *human behavior*, there is a place (as I think Wittgenstein would grant) for a science that makes *language* its object of study. This science would place seemingly meaningless phenomena regarding linguistic practice in a general framework that would enable us to *explain* why philosophical puzzlement/conceptual tensions occur, and enable us to *predict* that they will recur again under certain conditions (ways of interpreting language). I don't think taking philosophy to be a 'science' in this respect (a science of the misuses of language that lead to conceptual muddles) would have been unwelcome to Wittgenstein, as passages like the following suggest:

Language is an immense network of easily accessible wrong turnings. And so we watch one man after another walking down the same paths and **we know in**

**advance** where he will branch off, where walk straight on without noticing the side turnings, etc. etc. What I have to do then is erect signposts at all the junctions where there are wrong turnings so as to help people past the danger spots. [CV p. 18]

... it isn't as though everything we say has a conscious purpose; our tongues just keep going. Our thoughts run in established **routines**, we pass automatically from one thought to another according to the techniques we have learned. And now comes the time for us to survey what we have said. We have made a whole lot of movements that do not further our purpose, or that even impede it, and now we have to clarify our thought processes philosophically. [CV p. 64]

As long as there continues to be a verb 'to be' that looks as if it functions in the same way as 'to eat' and 'to drink', as long as we still have the adjectives 'identical', 'true', 'false', 'possible', as long as we continue to talk of a river of time, of an expanse of space, etc. etc., people **will keep stumbling over the same puzzling difficulties** and find themselves staring at something which no explanation seems capable of clearing up. [CV p. 15]

He [Wittgenstein] held that though the '**new subject**' must say a great deal about language, it was only necessary for it to deal with those points about language which have led, or are likely to lead, to definite philosophical puzzles or errors.<sup>133</sup>

The reason we 'know in advance' where 'someone will walk straight without noticing the side turnings' is that we have observed this phenomena *repeatedly*. Observed 'routines' enable us to frame generalizations and predictions. Just like the child who characteristically overgeneralizes and says 'sleped' instead of 'slept', 'brang' instead of 'brought',<sup>134</sup> philosophers too are bound to 'stumble over the same puzzling difficulties' by overstretching analogies. It is part of the philosopher's task to bring to conscious awareness the unacknowledged misuses of language (overstretched analogies) that lie behind these puzzling difficulties.

In the respects I have outlined, I don't think it would be wrongheaded to conclude that Wittgenstein *was* in the business of constructing 'theories'. Like the good scientist,

---

<sup>133</sup> Moore [1955, p. 27].

Wittgenstein collected a seemingly endless stream of empirical evidence (drawn from actual linguistic practice), which served as a basis for generalizations (about how certain features of language can mislead us), explanations (of how pseudo-problems arise), and predictions (about how they are bound to rise again and again, given certain features of language). A science of this sort is surely one that is desperately needed (efforts to understand the ‘human condition/predicament’, to borrow some post-modern jargon, demand it), and we can usefully see Wittgenstein’s philosophical method, applied to diverse regions of language, as inviting its establishment. If there is anything that unifies Wittgenstein’s philosophical explorations in these disparate linguistic settings, it is precisely his desire to get us to see how certain (deeply entrenched) features of language can lead us (and have led many philosophers) astray.

It is a great irony that Wittgenstein’s philosophical method, which has been condemned by many of his fiercest critics as being the *most belligerent* toward science, adheres more strictly to the canons of science as I have described them above than the approach of many of the staunchest defenders of a scientific approach to philosophical questions! One factor that obstructs this point from view is that Wittgenstein was mostly preoccupied with *gathering* empirical data, leaving the task of *organizing* it, and drawing the proper morals (generalizations, predictions) to the reader. As Wittgenstein cryptically puts this point:

If this stone won’t budge at present and is wedged in, move some of the other stones round it first. -- All we want to do is straighten you up on the track if your coach is crooked on the rails. Driving it afterwards is something we shall leave to you. [CV p. 39]

---

<sup>134</sup> These examples are offered by Chomsky [1986 p. 227].

What distinguished Wittgenstein's 'theories' from other philosophical theories (like *platonism* in the philosophy of math, or *psychologism* in the philosophy of mind) was that their *aim* was not to *solve* pseudo-problems, generated by overstretched analogies, but to *dissolve* them; to explain them away; to show why they are illegitimate. For Wittgenstein, there's nothing wrong with philosophical theorizing *per se*; there *is* something wrong with philosophical theorizing designed to solve pseudo-problems, as science does scientific problems. As he says, e.g.

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). [PI §125]

We can sum up Wittgenstein's advice to the reader as: 'Don't wed yourself to a "philosophical problem" too quickly! Don't become too dazzled by it! It might be a *snare!*'

Since pseudo-problems are dissolved and not solved, *no new knowledge* is gained.

As Moore recounts of Wittgenstein:

He also said that he was not trying to teach us any new facts: that he would only tell us trivial things – 'things which we all know already'; but that the difficult thing was to get a synopsis of these trivialities, and that our 'intellectual discomfort' can only be removed by a synopsis of *many* trivialities – that 'if we leave out any, we still have the feeling that something is wrong.' In this connection he said it was misleading to say that what we wanted was an 'analysis', since in science to 'analyze' water means to discover some new fact about it, e.g. that it is composed of oxygen and hydrogen, whereas in philosophy 'we know at the start all the facts we need to know.'<sup>135</sup>

The goal of ‘analysis’ in philosophy, as Wittgenstein conceived it, is to clarify the ‘grammar’ of our words, i.e., the appropriate contexts for using those words. This involves describing what we already know but have forgotten. There is of course a sense in which ‘analyses’ of this sort are *hidden* from view (and so we might think that analyzing here means finding something new), but that is only because we *overlook the familiar* aspects of word-usage, and need to be reminded of them.

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.) [PI §129]

How hard it is to see what is *right in front of my eyes!* [CV p. 39]

### 3.3.2 Exploring Imaginary Uses of Language

I turn now to the method of constructing hypothetical ‘language-games’. Wittgenstein gives an indication of what ‘language-games’ are in the *Blue Book*.

I shall in the future again and again draw your attention to what I shall call *language-games*. These are ways of using signs simpler than those in which we use the signs of our highly complicated everyday language. *Language-games* are the forms of language with which a child begins to make use of words. The study of *language-games* is the study of primitive forms of language or primitive languages.... When we look at such simple forms of language the mental mist which seems to enshroud our ordinary use of language disappears. We see activities, reactions, which are clear-cut and transparent. [BB p. 17]

---

<sup>135</sup> Moore [1955, p. 27].

The notion of a ‘language-game’, like that of a ‘conceptual framework’ or ‘discourse’ or ‘practice’, is a loose concept: it has ‘no precise definition or decisive and non-arbitrary criterion of individuation’.<sup>136</sup> Wittgenstein uses the notion of a ‘language-game’ to refer not only to *actual* uses of language – which might involve very basic ‘moves’, as in teaching language to children, or more sophisticated/complex ‘moves’, as we find in everyday discourse – but also to *hypothetical or invented* uses of language, which again might be basic or complex. **Hypothetical language-games** involve imaginary uses of language that are meant to be *compared* to actual language-games. They involve constructing ‘objects of comparison which are meant to throw light on the facts of our language by way of similarities and dissimilarities’. [PI §130] A hypothetical language-game is a method of instruction, and an indispensable one at that:

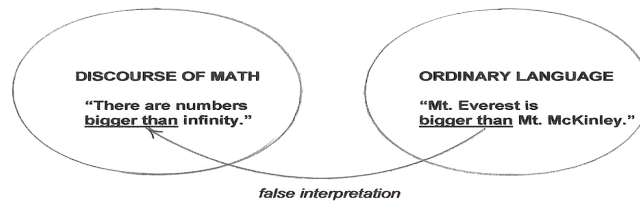
Nothing is more important for teaching us to understand the concepts we have than constructing fictitious ones. [CV p. 74]

What is involved in constructing a ‘fictitious concept’ is simply *assigning a new role* to a concept (one distinct from its actual role), *rearranging* the phenomena of language, so to speak, and then exploiting the dissimilarity between the invented and actual language-games to make a grammatical point: to bring out a new angle on the concept, to establish a new boundary on it, or bring down an old one. This is a subtler way of deconstructing pseudo-problems than the method discussed in the preceding section – of describing *actual* uses of language -- since it operates in an *indirect* manner. Hypothetical language

---

<sup>136</sup> Barnett’s expression [1990 p. 49].

games involve abstracting concepts from their normal circumstances of application, as in our earlier diagram.



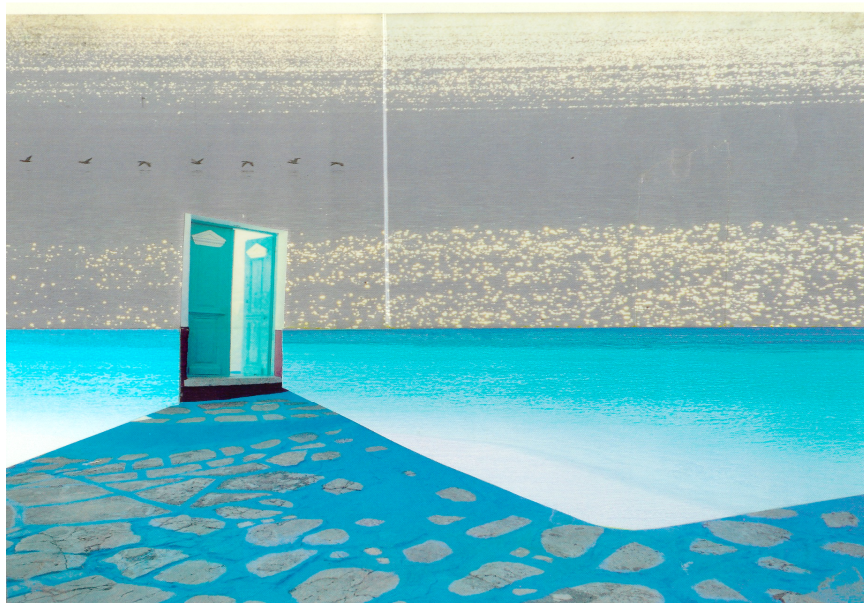
‘It dispenses the fog to study the phenomena of language in such imagined uses of language’, Wittgenstein remarks, ‘for here we can command a clear view of the aim and functioning of words’. [PI §5] Assembling reminders, rearranging facts, and assimilating pictures ‘[alter our way] of looking at things’ [PI §144]; they help put things into view.

One of the most important methods I use is to imagine a historical development for our ideas different from what actually occurred. If we do this we see the problem from a completely new angle. [CV p. 37]

But how does it ‘disperse the fog’ to take such departures from reality? How do language-games help us see things from a ‘completely new angle’?

In answering this, I have found it useful to draw an analogy with *surrealism* in art. One of the functions of art is to provide a *new interpretation* of our surrounding world, or to uncover some *hidden* aspect of it. Surrealism accomplishes this by shifting objects from their familiar/unique setting and placing them in an unfamiliar/alien context; typically one that forms a *contrast* with their original home. The contrast might be one involving a shift from darkness to light, old to new, noise to silence, etc. For example, a

surrealist work might involve moving a monastery from a craggy hill to the edge of the sea, or into a busy city, or some element from the busy city (say, the street with its pedestrians) onto a deserted island, etc. In this way the artist *draws our attention* to a given object -- an object we might have *overlooked* had it been in its familiar setting. He thereby sharpens our focus on it. An example to illustrate.<sup>137</sup>



This technique helps bring out ‘hidden’ aspects of the world, ‘hidden’ because they are so *familiar* that we overlook them; we take them for granted. They become insignificant from their familiarity. By *rearranging* phenomena, the artist makes the insignificant significant. Wittgenstein’s method of constructing hypothetical language-games can be viewed as an employment of a similar technique, only in a different medium: *language*. By taking us on journeys into ‘imaginary landscapes’ (imaginary uses of language),

---

<sup>137</sup> © Chrysoula Gitsoulis

where words/sentences have a function that differs from their actual function, Wittgenstein helps us to see the familiar role we have overlooked: how they *actually* function. For it is forgetting this that gives rise to philosophical puzzlement. Wittgenstein's method, as with surrealism, does not involve building a *new* construction out of new material, but only 'rearranging what we have always known' [PI §109], like the 'rearrangement of books in a library'. [BB p. 44]<sup>138</sup>

Unlike most surrealist works, however, Wittgenstein's re-arrangements of language are not *arbitrary*, but strategically chosen. Pseudo-philosophical problems, remember, are generated by overstressing the 'limits of language': interpreting the sense and function of words/sentences in one use-context/'language-game' in terms of their sense/function in another use-context/'language-game'. It is in this respect that we overstretch the 'limits of language', thereby generating *false interpretations* that lead to pseudo-philosophical problems. In a hypothetical language-game we are invited to imagine a world where the false interpretation *is* correct. *What else* would have to be true if it is correct? A hypothetical language-game fills in the details. The idea is that by accumulating enough of these details, we might at last come to see that we have been in the grip of a false interpretation. So hypothetical language-games, we can say, function like *reductio arguments*. This is why this aspect of Wittgenstein's method is more subtle than that of describing actual uses of language: it is an indirect method of 'proof' -- of getting us to see our mistaken assumptions.

An example to illustrate. In the note to PI §151, Wittgenstein tries to clarify the grammar of our concept of understanding. Can it be correct to think of understanding as

---

<sup>138</sup> Or the rearrangement of furniture in a room. According to Moore, Wittgenstein 'compared his method to the tidying of a room where you have to move the same object several times before you can get the room

a ‘mental state’? Well, suppose it *is* correct. What else would have to be true? To answer this, we need to consider how we *actually* employ the concept of a ‘mental state’. We apply it to experiences such as depression, excitement, and pain, among others. And what is true about how we apply these terms? We say, e.g., as Wittgenstein notes:

Actual use-context

‘He was depressed the whole day.’

‘He was in great excitement the whole day.’

‘He has been in continuous pain since yesterday.’

‘When did your pains get less?’

But can we speak of ‘understanding’ in these ways? To answer this, we need to construct imaginary use-contexts where the term *is* used in these ways. Again, using Wittgenstein’s examples:

Imaginary use-context

‘Since yesterday, I have understood this word.’ ‘Continuously’ though?

‘When did you stop understanding the word?’

As we can see, it is ‘ungrammatical’ to apply certain *temporal* concepts to ‘understanding’ in the way we do to mental states: we speak of being in continuous pain, or of a pain being interrupted for several minutes, or of suddenly ceasing to feel pain, but

it is grammatically jarring to regard understanding as ‘clockable’<sup>139</sup> in this way. Moreover, it is grammatically jarring to apply concepts of *intensity* to understanding in the way we do to mental states: we describe a pain as intense, or an emotion as strong, but it is ‘ungrammatical’ to describe understanding in these terms.<sup>140</sup> One might do so, but then he is using the term in nonstandard or conventionally unacceptable ways. Similarly, one might use a kitchen table as a TV stand, but this is not conventionally acceptable, and to bring this out to someone who decides to use it in this way, you might try placing a checkered table cloth over the table, or kitchen chairs around it, and a napkin holder next to the TV (you would be creating an imaginary language-game). If this provokes laughter, you have made your point, though indirectly of course. It is similar with words. Laughter might be the appropriate response to someone who uses words as in the imaginary use-context above (‘I stopped understanding the word yesterday’, etc.). Laughter is a *sign* that an interpretation is out of place, as in the joke:

Patient: I broke my arm in two places.

Doctor: Don’t go to those places.<sup>141</sup>

This is why Wittgenstein makes use of jokes to illuminate concepts. They are supposed to help us recognize that we are in the grip of a false interpretation.

Rearrangement of familiar facts in Wittgenstein’s hypothetical language-games is another reason (in addition to those mentioned in §3.1) Wittgenstein’s writing gives rise

---

<sup>139</sup> McGinn’s expression [1984 p. 5].

<sup>140</sup> We can also see Wittgenstein’s so-called ‘private language argument’ as taking the form of a reductio.

to a feeling of *eeriness* (a bit like the feeling we have as tourists traveling to unfamiliar lands). Witness for example the following language-games:

Now think of the following use of language: I send someone shopping. I give him a slip marked 'five red apples'. He takes the slip to the shopkeeper, who opens the drawer marked 'apples'; then he looks up the word 'red' in a table and finds a color sample opposite it; then he says the series of cardinal numbers – I assume that he knows them by heart – up to the word 'five' and for each number he takes an apple of the same color as the sample out of the drawer. – It is in this and similar ways that one operates with words. ... [PI §1]

Let us imagine a language for which the description given by Augustine is right. The language is meant to serve for communication between a builder A and an assistant B. A is building with building-stones: there are blocks, pillars, slabs and beams. B has to pass the stones, and that in the order in which A needs them. For this purpose they use a language consisting of the words 'block', 'pillar', 'slab', 'beam'. A calls them out; -- B brings the stone which he has learnt to bring at such and such a call. – Conceive this as a complete primitive language. [PI §2]

The eerie (or comic) aspect arises from using *familiar* words in *unfamiliar* ways, just as in surrealism it arises from rearranging elements of familiar settings. The shopping expedition would not, *in real life*, be conducted in the manner portrayed in PI §1. No grocer keeps apples in drawers labeled 'apples' or consults color charts. We do not live in such ludicrous, mechanical worlds. The point is, we are supposed to *contrast* that imaginary language-game with our actual linguistic practices. In this way we might see (the contrast might help bring to light) that communication does not demand that every word must have something for which it stands -- that something being its meaning -- as the *Augustinian picture* (a pseudo-picture) assumes.

What we have overlooked, what we have forgotten, for Wittgenstein, is often what is right before us: how language *actually* functions. It is to this world (for Wittgenstein

---

<sup>141</sup> This joke is from Richard Gilmore [1999, p. 96].

both the source of and final court of appeal for philosophical disputes) that Wittgenstein was constantly drawing our attention (either directly or indirectly), by making the insignificant *significant*, the ordinary *extraordinary*. It is perhaps in this respect more than any other – in recognizing the elementary sources of confusion that lie at the root of many of our seemingly most ‘profound’ philosophical problems -- that his work was revolutionary; and therein, I think, lies his most valuable contribution to philosophy.

## 4

## An Assessment of Kripke's Reading of Wittgenstein

---

 CONTENT
 

---

4.1	The Skeptical Problem of PI
4.2	The Skeptical Argument of PI
4.3	How False Interpretations Generate a Pseudo-Problem
4.4	The Skeptical Solution of PI: Dissolving the Pseudo-Problem
4.4.1	Dissolving the 'Inner State' Model of Meaning
4.4.2	Dissolving the Metaphysical Constraint in [J]
4.4.3	Dissolving the Epistemological Constraint in [J]
4.5	Conclusion

---

### 4.1 THE SKEPTICAL PROBLEM OF PI

In the last chapter, I tried to explain what Wittgenstein took pseudo-philosophical problems to be – that is, his means of identifying them – and his means of ‘treating’ them, by focusing primarily on his so-called ‘meta-philosophical’ remarks in §§107-136 of PI. These sections immediately precede his discussion of meaning and rule-following in §§137-242. As we saw, Wittgenstein likened the emergence of pseudo-philosophical problems to symptoms of a disease, not in need of *solution*, but of *dissolution*.<sup>142</sup>

---

<sup>142</sup> The important idea that Wittgenstein’s discussion of meaning can be properly understood only through the lens of his metaphilosophy was articulated in detail by Paul Horwich in his Fall 2001 CUNY graduate seminar, which I attended.

Pseudo-problems, as we saw, arise from false interpretations of language. In §§137-242, it is the problem of what understanding or meaning something by an expression ‘consists in’ [§§146, 148, 184, 186, 198] that is falsely interpreted by Wittgenstein’s interlocutor (who plays the role of the skeptic in WRPL). This leads him to generate a pseudo-problem, which, I submit, can be identified as the skeptic’s interpretation of the skeptical problem in WRPL. Thus, I try to show that the skeptical problem does emerge in PI, but as just another pseudo-problem, not in need of *solution*, but of *dissolution*. Throughout §§137-153 of PI, repeated attempts are made to solve this pseudo-problem, to no avail. It is mainly in these sections, I try to show, that we can locate the ‘skeptical argument’ of WRPL. In §§154-155, 179-183 of PI, Wittgenstein urges us to examine the kinds of circumstances in which we use the expressions that were falsely interpreted by the interlocutor (and led to the construction of a pseudo-problem), and the role these expressions have in our lives. It is in these sections, I try to show, that Wittgenstein presents the framework for a ‘skeptical solution’ -- a means of *dissolving* the pseudo-problem -- whose details he proceeds to fill in in §§184ff. Thus, the overarching aim of this chapter will be to show that the key ingredients of Kripke’s interpretation of Wittgenstein can be found in PI, and are adequately supported by it.

The pseudo-problem is generated because, like Kripke’s skeptic, the interlocutor reasons that any fact – let’s call it F – that understanding or meaning something by a word *w* ‘consists in’ must have a certain nature and role, which translates into a metaphysical and epistemological constraint, respectively, on F:

- (1) **Metaphysical constraint** (concerning the *normativity* of meaning): F contains a norm, a standard of correctness, which determines in advance how *w* is to be applied in an open-ended range of situations.
- (2) **Epistemological constraint**: I have direct (non-inferential) access to F, and I *consult* F with each fresh application of *w*.

In addition, the interlocutor reasons that F must be an internal ‘state’ of an agent; something which serves as the ‘source’ [§146] of meaning and understanding. That all these conditions be met is suggested in the reasons the interlocutor offers throughout, e.g., §138-153, for rejecting a number of candidates for what meaning something by an expression, or understanding it, consists in, and by the nature of the kinds of candidates considered.

I will refer to (1) and (2) as the **interlocutor’s justification constraint** [J] on any meaning-constituting fact. The main difference between the interlocutor’s [J] and the skeptic’s [J<sub>I+</sub>] is that [J] is sometimes characterized by the interlocutor in epistemological (reason-giving) terms, as above, and sometimes in causal/mechanistic terms. Causal readings are suggested in sections like 140, where the interlocutor is led to think that F (for example, an image coming before my mind) ‘forces’ a particular application on me, or ‘logically compels’ me to apply a word thus-and-so, or in §188, where he is led to think that my applications of a word are ‘in some unique way predetermined, anticipated’. They are also suggested in the ‘machine as symbol’ analogy [§§193, 194], where my applications of a word are compared to the future movements of an ‘ideally

rigid machine which *can* only move in such and such a way' [§194]. The machine's action 'seems to be in it from the start' [§193], 'in a mysterious way – already present' [§193], as are my applications of a word in my understanding or meaning something by it. Epistemological readings are suggested in §§211, 212, 213, 217, 219, and 228, where Wittgenstein rejects the interlocutor's assumption that we are always guided by, or can offer, *reasons* to justify our applications of words. In WRPL, Kripke also sometimes characterizes the justification constraint in causal terms, in particular when comparing the skeptical problem to Hume's 'skeptical problem' about causation. But throughout the skeptical argument, it is the epistemological characterization that dominates; this is why I formulated the skeptic's justification constraint epistemologically.

The problem that dominates §§138-153 is whether there is any fact that satisfies [J]. Wittgenstein runs through a number of candidates that might be taken to solve this 'pseudo-problem', and finds them lacking. The candidates considered and rejected include actual use [§§138,139,147], images coming before my mind [§§152,198,213], special feelings or experiences had when applying words [§§155, 179], rules coming before my mind [§§152,198,213], and dispositional states [§§146,149,150,153].

Let's look at the details.

## 4.2 THE SKEPTICAL ARGUMENT OF PI

### A. *Actual Use*

Actual use fails to satisfy [J], since I don't have *direct (non-inferential) knowledge* of how I actually use a word, given that it is *extended* in time.

[138] ... we *understand* the meaning of a word when we hear or say it; we grasp it in a flash, and what we grasp in this way is surely something different from the 'use' which is extended in time!

[139] When someone says the word "cube" to me, for example, I know what it means. But can the whole *use* of the word come before my mind, when I *understand* it in this way? ...

[147] "But how can it be? When *I* say I understand the rule of a series, I am surely not saying so because I have *found out* that up to now I have applied the algebraic formula in such-and-such a way! In my own case at all events I surely know that I mean such-and-such a series; it doesn't matter how far I have actually developed it." ...

Your idea, then, is that you know the application of the rule of the series quite apart from remembering actual applications to particular numbers. And you will perhaps say: "Of course! For the series is infinite and the bit of it that I can have developed finite."

### ***B. Contentless Introspectible Qualitative States of Consciousness***

Mental images fail to satisfy [J]: they fail to determine how I *ought* to apply a word in each fresh instance of application.

[139] ... What really comes before our mind when we *understand* a word? – Isn't it something like a picture? Can't it *be* a picture?

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.

...

The picture of the cube did indeed *suggest* a certain use to us, but it was possible for me to use it differently.

[139n] ... I see a picture; it represents an old man walking up a steep path leaning on a stick. – How? Might it not have looked just the same if he had been sliding downhill in that position? Perhaps a Martian would describe the picture so. ...

[140] ... 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.

[239] How is he to know what colour he is to pick out when he hears “red”? – Quite simple: he is to take the colour whose image occurs to him when he hears the word. – But how is he to know which colour it is ‘whose image occurs to him’? Is a further criterion needed for that? (There is indeed such a procedure as choosing the colour which occurs to one when one hears the word “....”) ...

It won’t help to bring to mind another picture for interpreting the original, for this just sets us off on a *regress*.

[141] Suppose, however, that not merely the picture of the cube, but also the method of projection comes before our mind? – How am I to imagine this? – Perhaps I see before me a schema shewing the method of projection: say a picture of two cubes connected by lines of projection. – But does this really get me any further? Can’t I now imagine different applications of this schema too? ...

Mental experiences fail to satisfy [J] because they too fail to determine how I *ought* to apply a word in each fresh instance of application.

[155] Thus what I wanted to say was: when he suddenly knew how to go on, when he understood the principle, then possibly he had a special experience – and if he is asked: ‘What was it? What took place when you suddenly grasped the principle?’ perhaps he will describe it much as we described it above – but for us it is the circumstances under which he had such an experience that justify him in saying in such a case that he understands, that he knows how to go on.

[179] ... one might think that the sentence “I can go on” meant “I have an experience which I know empirically to lead to the continuation of the series.” But does B mean that when he says he can go on? Does that sentence come to his mind, or is he ready to produce it in explanation of what he meant?

No. The words “Now I know how to go on” were correctly used when he thought of the formula: that is, given such circumstances as that he had learnt algebra, had used such formulae before. ...

### C. ‘Rules’ Coming Before the Mind

Rules coming before my mind fail to satisfy [J]: again, they fail to determine how I *ought* to apply a word in each fresh instance of application. Hence they too are ruled out as the constituents of meaning and understanding. As Colin McGinn puts it: ‘a sign is itself just a piece of lifeless syntax, and syntax can never add up to meaning; so it is hopeless to conceive of grasp of meaning as the association of signs. For no sign is *intrinsically* meaningful.’ [1984 p. 17]

[152] ... “B understands the principle of the series” surely doesn’t mean simply: the formula “ $a_n = \dots$ ” occurs to B. For it is perfectly imaginable that the formula should occur to him and that he should nevertheless not understand. “He understands” must have more in it than: the formula occurs to him. And equally, more than any of those more or less characteristic *accompaniments* or manifestations of understanding.

[198] “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. ...

It will not help to think of a rule coming before my mind as a kind of *inner intuition* – the idea being that the rule would be in a better position to guide or direct me if it ‘speaks to me’ or ‘tells me’ what to do. For this just sets us off, once again, on a *regress*:

[213] ... If intuition is an inner voice—how do I know *how* I am to obey it? And how do I know that it doesn’t mislead me? For if it can guide me right, it can also guide me wrong.

((Intuition an unnecessary shuffle.))

#### *D. Dispositional States*

If we take understanding to be a kind of know-how or ability,

[150] The grammar of the word ‘knows’ is evidently closely related to that of ‘can’, ‘is able to’. But also closely related to that of ‘understands’. (‘Mastery’ of a technique.)

[151] But there is also this use of the word “to know”: we say “Now I know!” – and similarly “Now I can do it” and “Now I understand!” ...

and we identify this know-how with dispositions to apply words, which we in turn identify with brain states,

[149] If one says that knowing the ABC is a state of the mind, one is thinking of a state of a mental apparatus (perhaps of the brain) by means of which we explain the *manifestations* of that knowledge. Such a state is called a disposition....

then we can conclude that linguistic dispositions fail to satisfy [J] because they too fail to determine how I *ought* to apply a word in each fresh instance of application.

[146] ... Perhaps you will say here: to have got the system (or, again, to understand it) can't consist in continuing the series up to *this* or *that* number: *that* is only applying one's understanding. The understanding itself is a state which is the *source* of the correct use.

What is one really thinking of here? Isn't one thinking of the derivation of a series from its algebraic formula? Or at least of something analogous? But this is where we were before. The point is, we can think of more than *one* application of an algebraic formula; and every type of application can in turn be formulated algebraically; but naturally this does not get us any further. ...

Moreover, conceived of as brain states, linguistic dispositions fail to satisfy the epistemological requirement in [J]: that I have *direct (non-inferential) knowledge* of F.

[153] We are trying to get hold of the mental process of understanding which seems to be hidden behind those coarser and therefore more readily visible accompaniments. But we do not succeed; or, rather, it does not get as far as a real attempt. For even supposing I had found something that happened in all those cases of understanding, -- why should *it* be the understanding? And how can the process of understanding have been hidden, when I said "Now I understand" *because* I understood?! And if I say it is hidden -- then how do I know what I have to look for? I am in a muddle.

Given that all candidate facts that understanding can seemingly be identified with -- be it something *extended* in time (such as actual use), or present in a *single instance* of time (such as an object coming 'before my mind' or a brain state) -- fail to satisfy [J], the interlocutor is led to believe that there is nothing that meaning 'consists in'.

It's important to note that [J] is not a condition that Wittgenstein himself insists be satisfied by any purported meaning-constituting fact; it is the interlocutor who insists on this. Why we will see shortly. Likewise, in WRPL, it is Kripke's skeptic, not KW, that

imposes the constraint [J<sub>I</sub>]. Kripke's skeptic plays the role of Wittgenstein's interlocutor, and KW the role of Wittgenstein.

### 4.3 HOW FALSE INTERPRETATIONS GENERATE A PSEUDO-PROBLEM

We describe competence in the use of an expression as a matter of 'grasping a rule' or 'following a rule' that 'determines in advance' how we ought to apply a word, and what sorts of application will conform to the rule. We take the rule to determine how a word ought to be applied in an indefinitely large range of cases, even though at any given time we have considered only a small fraction of these. We think of the rule as 'guiding' us in each new instance of application, and of someone's having learned the rule as that which 'explains' or is the 'source' of correct use. When we consider a new instance of application of a word, we confidently judge that 'the rule we have all along had in mind' dictates whether we ought to apply the word or not.

Like Kripke's skeptic, Wittgenstein's interlocutor *inflates* these seemingly harmless ways of speaking, thereby creating a mythological picture of what it is to understand or mean something by an expression, and a pseudo-problem about meaning and understanding. He comes to interpret grasp of a rule as 'the internalization of an open-ended set of preordained requirements, an informational state accessible ... only by a kind of guesswork.'<sup>143</sup> He pictures 'the relation between receiving an illustration of a

---

<sup>143</sup> Wright's expressions [2001 p. 164], though he doesn't apply them to the skeptical problem as I do here.

practice and going on to participate in it successfully as essentially mediated by cognition of the requirements of something which has been interiorized.’<sup>144</sup>

In §§186-197 (and in some earlier sections) of PI, Wittgenstein discusses the nature of the misconceptions of how language works that led to the generation of the pseudo-problem. I summarize the key points that can be extracted from that discussion below.

---

<sup>144</sup> Ibid p. 165



As we can see, by falsely interpreting expressions like those in the left-hand column, the interlocutor is led to believe that any fact F that understanding or meaning something by a word w ‘consists in’ must be a state having a certain kind of *nature* and *role*, which I summed up in [J]. This leads him to posit a pseudo-problem: Is there a state that satisfies [J]?<sup>145</sup>

#### 4.4 THE SKEPTICAL SOLUTION OF PI: DISSOLVING THE PSEUDO-PROBLEM

##### 4.4.1 Dissolving the ‘Internal State’ Model of Meaning

##### *Conflating the Grammatical Function of Meaning-Ascriptions with Descriptive Propositions*

---

<sup>145</sup> Wittgenstein doesn’t deny the propriety of using expressions like those in the left-hand column; he just cautions against falsely interpreting them, by abstracting them from their actual use-context.

[195] “But I don’t mean that what I do now (in grasping a sense) determines the future use *causally* and as a matter of experience, but in a *queer* way, the use itself is in some sense present.” – But of course it is, ‘in *some* sense’! Really, the only thing wrong with what you say is the expression ‘in a queer way’. The rest is all right; and the sentence only seems queer when one imagines a different language-game for it from the one in which we actually use it. ...

[197] “It’s as if we could grasp the whole use of a word in a flash.”— And that is just what we say we do. That is to say: we sometimes describe what we do in these words. But there is nothing astonishing, nothing queer, about what happens. ...

[194] “When we do philosophy we are like savages, primitive people, who hear the expressions of civilized men, put a false interpretation on them, and then draw the queerest conclusions from it.”  
[194]

In §§154-155, 179-183, Wittgenstein urges us to examine the kinds of circumstances in which we use the expressions (left-hand column above) that were falsely interpreted by the interlocutor, and the role these expressions have in our lives. This will help us see that it is by ‘crossing different pictures’ [§191], conflating the use of expressions in ‘different language-games’ [§195], that the interlocutor was led to construct a pseudo-problem. It will also help us see that the pseudo-problem is not in need of *solution*, but of *dissolution*.

The work of drawing our attention to how the falsely interpreted expressions function within their *actual* use-context (their natural surrounding / original home), where they do their work, is done in WRPL by the skeptical solution. According to Kripke [p. 73], Wittgenstein’s ‘skeptical solution’ to the ‘skeptical problem’ replaces the question of the *Tractatus*: ‘What must be the case for the sentence “Jones means so-and-so by w” to be true?’ by two others in PI:

- i. Under what conditions can we **justifiably assert** (or deny) that Jones means so-and-so by w?
- ii. What is the **role and utility** in our lives of our practice of asserting (or denying) such statements?

The skeptical solution ‘provides both conditions under which we are justified in attributing concepts to others and an account of the utility of this game in our lives.’ [p. 95] ‘We have to see under what circumstances attributions of meaning are made [what circumstances actually license such assertions] and what role these attributions play in

our lives'. [p. 86] By 'justification conditions' for meaning-ascriptions, Kripke means the 'language-game' in which meaning-ascriptions are rightfully employed. [p. 74] Justification conditions are grounded in social practices; not in the workings of the mind. We need to ask what role and utility meaning-ascriptions have in the language-game in which they are employed. When we do so, according to KW, we will see that *our hunt for the fugitive meaning-constituting fact was entirely misconceived*, for meaning-ascriptions do not have the function of *describing* the world as being a certain way, even though we attach notions of truth and falsity to them. They do not have the function of making *factual* claims about the world. The skeptic (Wittgenstein's interlocutor) was seduced by misleading pictures; it is the role of KW's skeptical solution to help him see them for what they are, by examining the role and function of attributions of meaning in our language. What role, then, do meaning-ascriptions have, if not one of *describing* the world? Kripke considers the case of '+' in answering this. By asserting that 'S means addition by "+"', we *express our confidence* in a subject: we label him as someone whose statements involving '+' ought to be relied upon. If I say of a grocer, for example, that he means *addition* by '+', I am saying that he can be trusted not to cheat you when you buy goods at his shop, and if I assert that he does not mean *addition* by '+', I am warning you against buying goods from his shop. It is in ways like this that meaning-ascriptions involving '+' have a *role* and *utility* in our lives.

Passages such as the following in PI, BB, and Z lend support to this interpretation:

[PI §180] *This is how these words are used.* It would be quite misleading, in this last case, for instance, to call the words a "**description** of a mental state". – One might rather call them a "**signal**"...

[PI §182] ... The criteria which we accept for ‘fitting’, ‘being able to’, ‘understanding’, are much more complicated than might appear at first sight. That is, the game with these words, their employment in the linguistic intercourse that is carried on by their means, is more involved – the role of these words in our language other -- than we are tempted to think. ...

[BB §66] ‘we are strongly inclined to use the metaphor of something being in a peculiar state for saying that something can behave in a certain way. And this ... metaphor is embodied in the expressions ‘He is capable of ...’, ‘He is able to multiply large numbers in his head’, ‘He can play chess’: in these sentences the verb is used in the present tense, suggesting that the phrases are descriptions of states which exist at the moment when we speak...’

[Z §16] ‘The mistake is to say that there is anything that meaning something consists in.’

[Z §608] ‘... an organism can come into being even out of something quite amorphous, as it were causelessly; and there is no reason why this should not really hold for our thoughts, and hence for our talking and writing.’

[Z §609] ‘It is thus perfectly possible that certain psychological phenomena cannot be investigated physiologically, because physiologically nothing corresponds to them.’<sup>146</sup>

Like KW, Wittgenstein insists that we look at the circumstances in which attributions of meaning are made, and the **role** these attributions play in our lives, to ‘resolve

---

<sup>146</sup> In these sections from *Zettel*, Wittgenstein appears to be denying that psychological facts **supervene** on physiological facts. McGinn addresses these passages in some detail [1984 pp. 112-6], and maintains that Wittgenstein is mistaken in holding this view. Following is my summary of McGinn’s argument:

- i. Suppose understanding does not supervene on a person’s internal physical constitution.
- ii. Suppose, moreover, that differences in understanding must issue in differences in use.
- iii. Then there are events which differ physically but which have no differentiating physical explanation. This is tantamount to the admission or claim that some physical events have no physical explanation.

Crispin Wright [2001 p.153-154] insightfully points out that this is a flawed argument.

... denying the supervenience claim would not commit Wittgenstein to the actuality of what McGinn describes. It would involve at most that nothing precludes its possibility. But in fact even that does not follow. What follows is only that, *for all the concept of understanding has to say about the matter*, there could be behavioral, ergo physical events whose explanation was not to be found in the *internal* physical state of the behaving subject. That may seem bizarre. But it falls well short of the claim that such events would be physically inexplicable. Nothing in physics rules out the idea that the movements of a body be explained not in terms of any internal change in it but as a direct result of external influences, for instance.

philosophical paradoxes' [§182]. It is the *circumstances/language-game* in which I use meaning-ascriptions that determine whether they are *rightfully* or *justifiably* employed.<sup>147</sup>

[154] But wait – if “Now I understand the principle” does not mean the same as “The formula ... occurs to me” (or “I say the formula”, “I write it down”, etc.) – does it follow from this that I employ the sentence “Now I understand ...” or “Now I can go on” as a **description** of a process occurring behind or side by side with that of saying the formula?

If there has to be anything ‘behind the utterance of the formula’ it is *particular circumstances, which justify me in saying I can go on* – when the formula occurs to me.

Try not to think of understanding as a ‘mental process’ at all. – For *that* is the expression which confuses you. But **ask yourself: in what sort of case, in what kind of circumstances, do we say, “Now I know how to go on,”** when, that is, the formula *has* occurred to me? ...

[155] Thus what I wanted to say was: when he suddenly knew how to go on, when he understood the principle, then possibly he had a special experience – and if he is asked: “What was it? What took place when you suddenly grasped the principle?” perhaps he will describe it much as we described it above – but for us **it is the circumstances under which he had such an experience that justify him in saying in such a case that he understands,** that he knows how to go on.

[179] Let us return to our case (151). It is clear that we should not say B had the **right** to say the words “Now I know how to go on”, just because he thought of the formula – unless **experience shewed that there was a connexion between thinking of the formula – saying it, writing it down – and actually continuing the series.** And obviously such a connexion does exist. ...

... The words “Now I know how to go on” were **correctly used** when he thought of the formula: that is, given **such circumstances as that he had learnt algebra, had used such formulae before,** -- But that does not mean that his statement is only short for a description of all **the circumstances which**

---

<sup>147</sup> The import of these passages appears to have been overlooked by Wright [2007 p. 8] and others who ascribe a **quietist** stance to Wittgenstein concerning standards of correctness. As we saw in chapter 2, there is a class of basic terms where, as Wright puts it, ‘there is no such underlying rationalizing knowledge enabling [our competent use of these terms]’ [2007 p. 18] But from this he draws a further conclusion: ‘A *fortiori* there is no metaphysical issue about the character of the facts it is knowledge of, with platonism and communitarianism presenting the horns of a dilemma. The knowledge *is* the competence. Or so I take Wittgenstein to be saying.’ [2007 p. 18] It is this conclusion which informs the quietist stance Wright ascribes to Wittgenstein. However, it doesn’t follow from Wright’s observation about basic cases that Wittgenstein rejected the view that language games determine what counts as a ‘correct’ move, or regarded it as incoherent.

**constitute the scene for our language-game.** – Think how we learn to use the expressions “Now I know how to go on”, “Now I can go on” and others; in what family of language-games we learn their use. ...<sup>148</sup>

[180] *This is how these words are used.* It would be quite misleading, in this last case, for instance, to call the words a “description of a mental state”. – One might rather call them a “signal”; and we judge whether it was **rightly** employed **by what he goes on to do.**

[182] ... In what kind of **circumstances** would it count as a **justification** to say “When I answered ‘yes’ I could do it, only now I can’t”? ...

#### 4.4.2 Dissolving the Metaphysical Constraint in [J]

##### *Conflating Different Senses of ‘Determination’*

Kripke claims [p. 26 n. 19] that Wittgenstein’s remarks on machines [§§193-195] express an explicit rejection of dispositional and causal accounts of meaning and following a rule.<sup>149</sup> This claim strikes me as being too strong. What they do express, I believe, is a rejection of *false pictures* of what dispositions associated with understanding consist in, i.e., of their *nature*. Kripke is correct in noticing that ‘the dispositional theory views the subject himself as a kind of machine, whose potential actions embody the function. So in this sense the dispositional theory and the idea of the machine-as-embodying-the-function are really one.’ [p. 35 n. 24] Wittgenstein wants to show this is a false picture: how I am disposed to apply a term does not ‘determine’ future action in the way the blueprint of a

---

<sup>148</sup> These passages, along with 183, suggest that Wittgenstein is not equating justification conditions with truth conditions; in Kripke’s words, Wittgenstein replaces truth conditions with justification conditions.

<sup>149</sup> Though he adds that his discussion ‘might contain somewhat more of [his own] argumentation in support of Wittgenstein rather than exposition of Wittgenstein’s own argument’ [p. 6 n. 7]

machine, the ‘machine-as-embodiment-of-the-function’ (a machine thought of as operating in ideal conditions, where there is no friction and so forth: ‘machine-as-symbol’, for short), determines future movement, or the way a formula determines its expansion. The machine-as-symbol’s action ‘seems to be in it from the start’ [§193], just as an algebraic formula’s expansion seems to be there in it from the start, ‘its movement seems to be already completely determined’ [§193], it seems like it ‘*can* only move in such-and-such a way’ [§194]. We don’t think of its movement as ‘empirically determined in advance’ [§193], but as ‘in a mysterious sense -- already *present*’ [§193].<sup>150</sup> By ‘crossing pictures’ -- conflating how the machine-as-symbol *determines* future action with how linguistic dispositions *determine* future action -- the interlocutor is led to believe that there is a ‘super strong connexion’ [§197] between how I am disposed to apply a word and my future applications of it in an open-ended range of situations. He builds this false interpretation into [J]. Understanding, he is led to think, ‘forces’ a particular application on me, or ‘logically compels’ [§140] me to apply a word thus-and-so in any given situation, just as the machine-as-symbol determines in advance how the machine will behave in any given situation, or an algebraic formula determines in advance a given output for any input.

---

<sup>150</sup> It might seem mysterious that the machine-as-symbol can *contain* all its future movements, for unlike the movements of a physical machine which present physical conditions allow us to say *may* occur, the possibilities of movement of the machine-as-symbol are *already* present in some ‘queer’, ‘more determinate’ way. The mystery dissipates once we recognize that we are brought by our education (training) to use the picture of a machine as ‘the first of a series of pictures which we have learnt to derive from this one.’ [193] In my training, I am taught to take the picture as one in a series of pictures of the machine’s various states. I am taught that I am allowed to replace one picture with another, thus representing changing states in the machine, only in accordance with certain rules. ‘Suppose now that, having received this training, I imagine someone moving one of the machine’s parts. How I then suppose the other parts will dispose themselves in response to the movement will not be something I *speculate* about: it will simply follow from my having been trained to make only certain transitions among the machine’s pictured states. And this is all we need to explain how the machine’s possibilities of movements can be “in it from the start”: whatever manipulation I imagine, how the other parts accommodate

To dismantle this false interpretation, Wittgenstein draws our attention to the operation of concrete physical machines, which are also said to ‘determine’ a future course of action. But here, it is apparent that future movement is not logically guaranteed: their parts might bend, break off, melt, and so on. [§193] Noticing this might help us see that there is not a ‘super-strong connection’ [§197] between understanding and use, as the interlocutor thinks. In mastering or acquiring a term, a subject acquires a disposition of a particular sort. But his having acquired and continuing to possess the disposition does not mean that he will manifest it on *every* occasion in which its manifestation is called for: on some occasions, the manifestation of the disposition may be blocked or inhibited. I might, for example, mean *horse* by ‘horse’, yet be disposed to apply ‘horse’ to horsy-looking cows on a dark night. I might mean *addition* by ‘+’, yet fail to carry digits that end in 7’s when intoxicated. More generally, an object could possess a dispositional characteristic, and yet the manifestation of that characteristic be blocked or inhibited by some feature of the object’s situation. So how I am disposed to apply a term does not ‘force’ a particular application on me.

Moreover, and more importantly, as I pointed out in chapter 2, the function corresponding to how we are disposed to apply words, both on an individual and communal level is, for most words of our language, one that is **progressively constructed/created** over time, and hence not fully determined in advance, as the function corresponding to the behavior of the machine-as-symbol is. It does not reach ahead of us and determine on its own an output for any given input (use-context). We know, by contrast, how the machine-as-symbol will behave for any given input. In this

---

themselves to it comes to me *unhesitatingly*.’ [Falk 1994 p. 392] Ditto, as Wittgenstein explains in [PI 189], for the case of an algebraic formula *containing* its expansion.

sense, its behavior is *logically determined in advance*. But our linguistic behavior is not pre-determined in this way: for most words of our language, we can't determine in advance how we will apply them in any given use-context. As I pointed out in chapter 2, most words of our language, like the concept of a game, have open, not closed, boundaries (as we find in most of logic and mathematics). They are surrounded by 'grey areas' -- cases for which there are no rules of language that settle whether the term applies or not. In elucidating them, certain **models/paradigms** are usually given over which there can be no doubt or hesitation as to whether the word applies or not, but conditions of application cannot be *exhaustively* enumerated.<sup>151</sup> There will always be new situations that will require a **decision/judgment** on our part to either extend the concept to cover the new case or exclude it from its extension. In forming such

---

<sup>151</sup> The temptation to think otherwise arises, as Crispin Wright observes in his explication of Wittgenstein,

‘from the reflection that the explanation of a rule must eventually culminate in, or anyway ultimately be founded upon, the giving of illustrations of its application; and that any such illustrations are finite, and hence open in principle to an indefinite variety of interpretations. Yet explanations do usually, or so we suppose, secure mutual understanding. So somehow more is got across – the thought continues – than the pursuit of explanations can ever make completely explicit. Correct uptake of an explanation is having the right ‘something’ come into one’s mind as a result of the explanation; and the resulting informational state, though it is expressed in one’s subsequent practice with the concept concerned, essentially transcends it.’ [2001 p. 159]

The test for whether one has grasped the explanation depends on one’s ability to ‘go on’ to classify other things, not in the samples offered, in the ‘same way’. This may tempt one to describe the process of correct uptake as one that involves the formation of a rule that tells one how to go on in unexemplified cases. Wittgenstein, as Wright notes, is concerned to expose this mistaken epistemological picture, and in turn ‘the rules as rails imagery which imply a figurative expression of Platonism’. [2001 p. 160] ‘The connection between the training and explanations which we received and our subsequent practices ... is not mediated by the internalization of explanation-transcendent rules that, in our training, we (something like) guessed’. [2001 p. 167] ‘It is, for epistemological purposes, a *basic* fact about us that ordinary forms of explanation and training do succeed in perpetuating practices of various kinds – that there is a shared uptake, a disposition to concur in novel judgments involving the concepts in question. The mythology of the “rules as rails” attempts an explanation of this fact. But the truth is the other way round: it is the basic disposition to agreement which sustains all rules and rule-governed institutions. The requirements which our rules impose on us would not be violated if there were not this basic agreement; they would not so much as *exist*.’ [2007 p. 7]

decisions, we need to consider whether the new case is similar in enough important respects to the paradigm.

I take Wittgenstein to be making a similar point in passages like the following:

[186] “What you are saying, then, comes to this: a new insight – intuition – is needed at every step to carry out the order ‘+n’ correctly” – to carry it out correctly! How is it decided what is the right step to take at any particular stage? – “The right step is the one that accords with the order – as it was meant.” – So when you gave the order +2 you meant that he was to write 1002 after 1000 – and did you also mean that he should write 1868 after 1866, and 10036 after 10034, and so on – an infinite number of such propositions? – “No: what I meant was, that he should write the next but one number after *every* number that he wrote; and from this all those propositions follow in turn.” – But that is just what is in question: what, at any stage, does follow from that sentence. Or, again, what, at any stage we are to call “being in accord” with that sentence (and with the *meaning* you then put into the sentence – whatever that they may have consisted in). It would almost be more correct to say, not that an intuition was needed at every stage, but that a new **decision** was needed at every stage.<sup>152</sup>

As Crispin Wright points out, the point of the contrast between ‘intuition’ and ‘decision’ is that ‘the former implies and the latter repudiates the suggestion that – even in the most basic cases, where one can say nothing by way of justification for one’s particular way of proceeding – rule-following is a cognitive accomplishment, success in tracking an independently constituted requirement.’ [2001 p. 161]<sup>153</sup> This way of conceiving the

---

<sup>152</sup> See also PI §213, RFM VI, 24, Brown Book §5.

<sup>153</sup> Decision talk, unlike intuition talk (involved in apprehending latent essences), repudiates the suggestion that rule-following requires a ‘tracking’ (Crispin Wright’s term) epistemology; that we follow, as Wright puts it, ‘rules-as-rails, rules which somehow reach ahead of us and determine of themselves their every actual and counterfactual proper application’. [2001 p. 163] The rules-as-rails picture leads to insuperable difficulties: ‘We have no accountable idea of what would constitute the direction taken by the rule off its own bat, as it were... once the direction is conceived, after the fashion of Platonism, as determined autonomously, and our performance, whether communal or solitary, as merely an exercise in tracking’. [2001, p. 161] Wittgenstein was concerned to dismantle this rule-as-rails picture as it emerges in, e.g., logic and mathematics. The picture is one which, as Wright [2007 p. 3] describes it ‘views logic and mathematics as tracking absolutely hard conceptual structures and interconnections. Discoveries in mathematics are regarded as the unpacking of (in the best case) deep but (always) predeterminate implications of the architecture of our understanding of basic mathematical concepts, as codified in

*nature* of the dispositions associated with a subject's use of a word does not oppose the idea of future determination, but rather the mythical conception of what that determination consists in, as encapsulated in [J]; in particular, it opposes the idea that in some queer way *all* future applications are already present in the meaning-constituting fact.

#### 4.4.3 Dissolving the Epistemological Constraint in [J]

##### *Conflating Knowledge-How with Knowledge-That*

In the grip of a *false picture* of what linguistic competence consists in, the interlocutor reasons that I consult rules for guidance with each fresh application of *w*, as summed up in the *epistemological* constraint in [J]. Wittgenstein seeks to dislodge this constraint, which he sees as distorting the epistemology of rule-following. To begin with, how I am disposed to apply a word is habitual, immediate, unreflective, and automatic.

[212] When someone whom I am afraid of orders me to continue the series, **I act quickly, with perfect certainty**, and the lack of reasons does not trouble me.

[228] “We see a series in just *one* way!”—All right, but what is that way? Clearly we see it algebraically, and as a segment of an expansion. Or is there more in it than that?—” 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

---

intuitively apprehended axioms. And logical inference, for its part, as the tracing of steps which are, in some sense, -- in a favorite image of his -- *already* drawn and which we have no rational option but to acknowledge once presented to us. This way of thinking – it is, of course, Frege's way of thinking – conceives of the requirements of at least logical and mathematical rules as hyper-objective: as somehow constituted quite independently of any propensities for judgment or reaction of ours. So an account seems needed of *how* they are constituted and of how we might reasonably presume ourselves capable of keeping intellectual track of their requirements so conceived.’

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.**

Use of the notion of guidance involves an overly ‘intellectual’ conception of the phenomena of understanding – ‘it underestimates the primitiveness of what is going on’, as Colin McGinn puts it [1984 p. 16]. It is not that a rule is the *source* of a subject’s use of a word, in the manner that [J] requires; rather, exactly the opposite holds: how a subject uses a word determines what rule he is following, as Wittgenstein points out in passages like the following:

[146] ... Perhaps you will say here: to have got the system (or, again, to understand it) can’t consist in continuing the series up to *this* or *that* number: *that* is only applying one’s understanding. The understanding itself is a state which is the *source* of the correct use.

What is one really thinking of here? Isn’t one thinking of the derivation of a series from its algebraic formula? Or at least of something analogous? -- But this is where we were before. The point is, we can think of more than *one* application of an algebraic formula; and every type of application can in turn be formulated algebraically; but naturally this does not get us any further. -- The **application is still the criterion of understanding.**

But let’s suppose understanding is a matter of something, e.g., a rule, coming ‘before my mind’. How does the rule guide me or tell me what to do in novel cases? I might have a rule come ‘before my mind’ without knowing what it means. So it looks like I need to *interpret* it to figure out what its requirements are. But that can be done in lots of ways. So it looks like invoking a rule is of help only if I interpret it *correctly*. But now, as Crispin Wright observes, ‘To describe someone as “knowing the correct interpretation of the rule for the nth place” just becomes a piece patter equivalent to saying that he knows how to apply the rule at the nth place.’ [2001 p. 163] This is why Wittgenstein says:

[198] “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.** ...

[201] ... 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.

Moreover, in many cases, I can’t even produce a rule to justify a given application of a word. So (in addition to the earlier *phenomenological* point made above) we have reason to doubt that there are rules guiding all our applications of words, as [J] demands. I often lack grounds (reasons) for my applications in any useful general sense of the term: I just find myself using language, instinctively.

In some cases, I might be able to produce a rule of the kind the interlocutor demands. For example, if asked to justify my responding ‘125’ when asked the sum of 68 and 57, I might say that I added 8 and 7 to get 15, that I put down 5 and carried 1, and so on. But ultimately, I am bound to reach a level where I cannot justify my application of a term by *producing* an informative rule that ‘tells me’ how to apply the term, or, still less, *recognize* some informative expression as the rule I follow. Explanations cannot go on forever. I will reach a level of **basic terms** where I act without any reason in terms of which I can justify my action, where I act unhesitatingly but ‘blindly’; a level of ‘brute understanding’, as we might put it. Wittgenstein brings this point out in passages like the following:

[211] 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.**

[217] “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.”**

(Remember that we sometimes demand definitions for the sake not of their content, but of their form. Our requirement is an architectural one; the definition a kind of ornamental coping that supports nothing.)

[219] ... When I obey a rule, I do not choose. I obey the rule *blindly*.

[236] Calculating prodigies who get the right answer but **cannot say how**. Are we to say that they do not calculate? (A family of cases.)

Terms for directly sensed properties, such as ‘red’, ‘blue’, ‘hot’, ‘cold’, ‘sweet’, ‘bitter’, ‘harmonious’, ‘dissonant’, or evaluative terms like ‘beautiful’, ‘good’, ‘just’, ‘right’ aptly illustrate these remarks (as I pointed out in chapter 2). I might not be able to produce an informative rule that ‘tells me’ how to apply these terms. In the latter case, a certain object or event might create a certain feeling or mood in me (a confident inclination that this way is the *right* way to apply the term rather than another way), an unhesitating disposition to proceed in a particular way, which I cannot capture in an informative descriptive phrase. In the former case, were I pressed for a reason, I might say it looks like *that*, and *point* to the color of a red object (ditto for other terms from that category -- feels like *that*, sounds like *that*, tastes like *that*, smells like *that*). We will look in vain for any more substantial ‘explanation’. My understanding here (‘grasp of a rule’) is best construed as consisting in **knowledge how** to use them as opposed to **knowledge that** they mean such-and-such (*propositional knowledge*), or **explicit knowledge** of how to

apply them – knowledge that can be elicited with suitable enquiry or prompting (we leave it open how much prompting is allowable).

Wittgenstein compares this know-how to a skill or ability.

[150] The grammar of the word ‘knows’ is evidently closely related to that of ‘**can**’, ‘**is able to**’. But also closely related to that of ‘understands’. (**‘Mastery of a technique**)

[151] But there is also this use of the word “to know”: we say “Now I know!” – and similarly “Now I **can** do it” and “Now I understand!” ...

[199] ... To understand a sentence is to understand a language. To understand a language means to be **master of a technique**.

Viewing understanding in this way is (as I pointed out in chapter 2) in keeping with studies of the psychology of classification that, as Jackman [2003 p. 34] or Barbiero [\*\*\*\*] point out, suggest that we conceptualize experience in terms of *prototypes* rather than in terms of categories determined by sets of necessary and sufficient conditions. As Wittgenstein himself puts this point in PI §201, ‘There is a way of grasping a rule which is not an interpretation, but which is exhibited in what we call “obeying a rule” and “going against it” in actual cases’. Wittgenstein’s well known example of a **game** illustrates this point: the ways in which we employ this term may have no more unity than a rope has a single fiber, so that all I could do is point to *examples* of games – board games, card games, ball-games, Olympic games, etc. -- as opposed to a set of necessary and sufficient conditions (some exhaustive set of properties) that characterizes what all games supposedly share in common. ‘Don’t say: there *must* be something common, or they would not be called “games”’, Wittgenstein says in PI §66, ‘but *look and see* whether there is anything common to all. For if you look at them you will not see

something that is common to *all*, but similarities, relationships, and a whole series of them at that...' Not all games are amusing, nor is there always winning or losing or competition. Not all games require more than one player. Different parts are played by skill and luck. What we find is a 'complicated network of similarities overlapping and crisscrossing', such that we can say games form a family with '**family resemblances**' [PI §67] – threads of common features running among the many things we call 'games' -- but have no single common trait. Asked to justify my use of the term 'game', all I may be able to do is point to examples and say 'This and *similar* things are called games'. Understanding here is again best conceived as knowing *how* to use the term, which is exhibited in appropriate judgments about particular cases, as opposed to apprehending some manifest or latent essence (what all games supposedly have in common), as demanded by [J]

Wittgenstein equates the know-how involved in language mastery with **dispositions** to apply words in appropriate ways.

[187] ... When you said "I already knew at the time..." that meant something like: "if I had then been asked what number should be written after 1000, I should have replied '1002'." And that I don't doubt. This assumption is rather of the same kind as: "If he had fallen into the water then, I should have jumped in after him". ...<sup>154</sup>

The appropriate dispositions are acquired through **training** [§§5,6,189, 198, 206, 208].

---

<sup>154</sup> In PI §151n Wittgenstein refers to 'understanding a word' as a 'state'. This does not necessarily commit him to its being a *physical state* of my brain. As Colin McGinn points out [1984], what Wittgenstein wishes to record by calling understanding a 'state' here is that it is not an *event* in consciousness but an *abiding condition*, as we take a disposition to be.

[189] ... We use the expression: “The steps are determined by the formula”. *How* is it used?—We may perhaps refer to the fact that people are brought by **their education (training)** so to use the formula  $y = x^2$ , that they all work out the same value for  $y$  when they substitute the same number for  $x$ . Or we may say: “These people are so trained that they all take the same step at the same point when they receive the order ‘add 3’”. We might express this by saying: for these people the order “add 3” completely determines every step from one number to the next. (In contrast with other people who do not know what they are to do on receiving this order, or who react to it with perfect certainty, but each one in a different way.) ...

[190] It may now be said: “The way the formula is meant determines which steps are to be taken”. What is the criterion for the way the formula is meant? It is, for example, the kind of way we always use it, the way we are **taught** to use it. ...

[198] ... 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. ...

[206] Following a rule is analogous to obeying an order. We are **trained** to do so; we react to an order in a particular way. ...

Given my training, I have acquired a disposition to apply a word thus-and-so, though with a different training I could have acquired a different disposition. In so far as an individual is a member of a linguistic community, communal practices kick in to help bring his dispositions in line with the community’s. Communal practices include tendencies to correct, criticize, and adjust deviations. They serve as ‘normative model[s], the function of which is to constrain or regulate behavior’.<sup>155</sup> Mutual pressures toward conformity, which form a basic part of the training process, consisting of reinforcements and punishments, help bring one’s dispositions in line with correct patterns of usage. Individuals, in virtue of their membership in a linguistic community, internalize these pressures and join in requiring these patterns of behavior in themselves and others. It is a

---

<sup>155</sup> Barbiero’s expression [\*\*\*\*].

basic fact about us that ordinary forms of explanation and training do succeed in perpetuating ‘a shared uptake, a disposition to concur in novel judgments involving the concepts in question’.<sup>156</sup>

#### 4.5 CONCLUSION

In chapter 1, I presented a skeptical problem about meaning and understanding as developed by Kripke in WRPL which, I argued here, also emerges in the rule-following sections of PI. I argued, moreover, that both the skeptic (who plays the role of Wittgenstein’s interlocutor) and the interlocutor place a false interpretation on this problem, thereby creating a pseudo-problem. Finally, I argued that by describing actual linguistic practice, both Kripke’s Wittgenstein and Wittgenstein try to show that the pseudo-problem is not in need of solution, but of dissolution. I therefore conclude that WRPL can be seen as presenting a fair portrayal of the contested sections of PI, and not an alien one. To establish this conclusion, I felt it imperative to situate the skeptical problem, as interpreted by the skeptic, and the skeptical solution, into a larger framework concerning Wittgenstein’s so-called ‘metaphilosophy’ (remarks concerning the practice of philosophy), that deals with his identification and treatment of pseudo-philosophical problems. This would enable us to see the pseudo-skeptical-problem as taking its place among an assortment of pseudo-problems which occupied Wittgenstein’s attention, and which he sought to dissolve through the method I detailed in chapter 3. Recognizing this,

---

<sup>156</sup> Wright’s expression [2001 p. 167].

it is my hope, might help us develop a greater appreciation for the fundamental importance of his philosophical method in his rule-following discussion, and in his philosophy more generally.

## Bibliography

### Armstrong, David

[1996] 'Categoricalist versus Dispositionalist Accounts of Properties', *Acta Analytica*, 15, pp. 7-19.

[1992] (edited with Ullin T. Place) 'A Debate on Dispositions, Their Nature and Their Role in Causation – Part I – The Armstrong-Place Debate', *Conceptus*, 26(66), pp. 3-44.

### Baker, G & Hacker, P

[1993] *Wittgenstein: Meaning and Understanding, vol. I*, Oxford: The Clarendon Press. [Chapter 13]

[1984] [a] 'On Misunderstanding Wittgenstein', *Synthese*, 58, pp. 407-50.

[1984] [b] *Language, Sense, and Nonsense*, Oxford: Blackwell.

### Baker, Gordon

[1984] 'Critical Study: On Misunderstanding Wittgenstein: Kripke's Private Language Argument', *Synthese*, 58, pp. 407-450.

### Barbiero, Daniel

[\*\*\*\*] 'Chosmky V. Kripke: Round Two: Methodological Collectivism and Explanatory Adequacy'. [http://www.personal.kent.edu/~pbohanbr/Webpage/New/HuenBarbiero/kripke2\(Barbiero\).html](http://www.personal.kent.edu/~pbohanbr/Webpage/New/HuenBarbiero/kripke2(Barbiero).html).

### Barnett, David

[1990] 'The Rhetoric of Grammar: Understanding Wittgenstein's Method' *Metaphilosophy*, 21(1&2).

Blackburn, Simon

- [1998] 'Realism and Truth: Wittgenstein, Wright, Rorty and Minimalism', *Mind*, 107(425), pp. 157-181.
- [1993] 'Wittgenstein and Minimalism' in B. Garrett and K. Hulligan (eds) *ANU Working Papers in Philosophy, 4: Themes from Wittgenstein*, Canberra: Philosophy Program, Research School of Social Sciences, ANU, pp. 1-14.
- [1990] 'Wittgenstein's Irrealism' in R. Haller and J. Brandl (eds) *Proceedings of the International Wittgenstein Symposium*, 14, Vienna: Verlag Holder-Pichler-Temsky, pp. 13-26.
- [1984] [a] 'The Individual Strikes Back', *Synthese*, 58, pp. 281-301.
- [1984] [b] *Spreading the Word*, Oxford: Oxford Press. [Part I]

Boghossian, Paul

- [1993] 'Sense, Reference, and Rule-Following' in Villanueva, Enrique (ed) *Naturalism and Normativity*, Ridgeview: Astacardo.
- [1990] [a] 'The Status of Content', *The Philosophical Review*, XCIX(2), pp. 157-184.
- [1990] [b] 'The Status of Content Revisited', *Pacific Philosophical Quarterly*.
- [1989] [a] 'The Rule Following Considerations', *Mind*, 98, pp. 507-49.
- [1989] [b] 'Review of Colin McGinn's *Wittgenstein on Meaning*', *Philosophical Review*, 98, pp. 83-92.

Bouveresse, Jacques

- [1955] *Wittgenstein Reads Freud*, Princeton University Press.

Burge, Tyler

[1979] 'Individualism and the Mental' in P. French, T. Uehling, and H. Wettstein (eds) *Studies in Metaphysics, Midwest Studies in Philosophy*, 6, Minneapolis: University of Minnesota Press, pp. 73-122.

Byrne, Alex

[1996] 'On Misunderstanding Kripke's Wittgenstein', *Philosophy and Phenomenological Research* LVI(2).

Carnap, Rudolf

[1967] Autobiography, reprinted in *K. T. Fann (ed), Ludwig Wittgenstein: The Man and his Philosophy*, Harvester Press, 1967.

Cavell Stanley

[1989] *This New Yet Unapproachable America: Lectures After Emerson and Wittgenstein*, Albuquerque, N.M.: Living Batch Press.

Chomsky, Noam

[1986] *Knowledge of Language*, New York: Praeger. [Chapter 4]

Collins, Arthur

[1992] 'On the Paradox Kripke Finds in Wittgenstein', *Midwest Studies in Philosophy*, XVII.

Crittenden, Charles

[1970] 'Wittgenstein on Philosophical Therapy and Understanding', *International Philosophical Quarterly*, 10, pp 20-43.

Divers, John & Miller, Alexander

[1994] 'Best Opinion, Intention Detecting, and Analytic Functionalism', *Philosophical Quarterly*, 44, pp. 239-45.

Dummett, Michael

[1991] *The Logical Basis of Metaphysics*, Cambridge MA: Harvard University Press.

Ebbs, Gary

[1997] *Rule-Following and Realism*, Cambridge: Harvard University Press. [Chapters 1-3]

Edwards, Jim

[1998] 'Response-Dependence, Kripke, and Minimal Truth', *European Review of Philosophy*, 3, pp. 149-74.

[1992] 'Best Opinion and Intentional States', *Philosophical Quarterly* 42, pp. 21-33.

Engelmann, Paul

[1967] L. Furtmuller (trans), B. F. McGuinness (ed), *Letters from Ludwig Wittgenstein to Paul Engelmann with a Memoir*, Oxford: Blackwell, 1967.

Falk, Barrie

[1994] 'Doing what One Meant to Do', *Synthese* 98, pp. 379-399.

Forbes, Graeme

[1984] 'Skepticism and Semantic Knowledge', *Proceedings of the Aristotelian Society*, 84, *Supplementary Volume*, pp. 221-37.

George, Alexander

[1989] (ed) *Reflections on Chomsky*, Oxford: Blackwell.

Gill, Jerry

[1971] 'Wittgenstein and the Function of Philosophy', *Metaphilosophy*, 2(2).

Gilmore, Richard

[1999] *Philosophical Health: Wittgenstein's Method in 'Philosophical Investigations'*, Lexington Books.

Ginet, Carl

[1992] 'The Dispositionalist Solution to Wittgenstein's Problem about Understanding a Rule', *Midwest Studies in Philosophy*, 17, pp. 54-73.

Goldfarb, Warren

[1989] 'Wittgenstein, Mind and Scientism', *Journal of Philosophy*, 86.

[1985] 'Kripke on Wittgenstein on Rules', *Journal of Philosophy*, 82, pp. 471-88.

[1983] 'I Want You to Bring Me a Slab' *Synthese* LV1(3), pp. 265-282.

Guttenplan, Samuel

[1975] (ed) *Mind and Language*, Oxford: Oxford University Press.

Hacker, Peter

[1972] *Insight and Illusion*, Oxford: The Clarendon Press.

Hale, Bob

[1998] 'Rule Following, Objectivity and Meaning' in B. Hale and C. Wright (eds) *A Companion to the Philosophy of Language*, Oxford: Blackwell.

Haukioja, Jussi

[2002] 'Soames and Zalabardo on Kripke's Wittgenstein', *Grazer Philosophisch Studien*, 64, pp. 157-173.

Hilmy, Stephen S.

[1987] *The Later Wittgenstein: The Emergence of a New Philosophical Method*, Oxford: Basil Blackwell.

Holton, Richard

[1993] 'Intention Detecting', *Philosophical Quarterly*, 43, pp. 298-318.

[1992] 'Response-Dependence and Infallibility', *Analysis*, 52, pp. 180-4.

[1991] 'Intentions, Response Dependence, and Immunity from Error' in P. Menzies (ed) *ANU Working Papers in Philosophy #1: Response Dependent Concepts*, Canberra: Philosophy Program, Research School of Social Sciences, ANU, pp. 83-121.

Holtzman, S. & Leich, C.

[1981] (eds) *Wittgenstein: To Follow a Rule*, London: Routledge & Kegan Paul.

Horwich, Paul

[2005] *A Deflationary Point of View*, Oxford. [Chapter 10]

[1998] *Meaning*, Oxford: Oxford Press. [Chapters 8 & 10]

[1995] 'Meaning, Use and Truth', *Mind*, 104.

[1994] 'Meaning and Metaphilosophy', *Philosophy and Phenomenological Research*, LIV(1).

[1990] 'Wittgenstein and Kripke on the Nature of Meaning', *Mind and Language*, 5(2).

[1984] 'Critical Notice: Saul Kripke: Wittgenstein on Rules and Private Language', *Philosophy of Science*, 51, pp. 163-171.

Jackman, Henry

[2003] 'Foundationalism, Coherentism, and Rule-Following Skepticism', *International Journal of Philosophical Studies*, 11(1), pp. 25-41

Janik, A and Toulmin, S

[1973] *Wittgenstein's Vienna*, Simon and Schuster.

Katz, Jerrold

[1990] *The Metaphysics of Meaning*, MIT Press. [Chapter 4]

Kenny, Anthony

[1982] 'Wittgenstein on the Nature of Philosophy' in Brian McGuinness (ed),  
*Wittgenstein and his Times*, England: Thoemmes Press, pp. 1-26.

Kripke, Saul

[1984] *Wittgenstein on Rules and Private Language*, Harvard

Levin, Michael

[2007] 'Bundling Hume with Kripkenstein', *Synthese* 55(1), pp. 35-65.

Malcolm, Norman

[1962] *Ludwig Wittgenstein: A Memoir*, London: Oxford Press.

Martin, C. B. & Heil, John

[1998] 'Rules and Powers' in James Tomberlin (ed) *Language, Mind and Ontology*,  
Philosophical Perspectives, 12, Cambridge: Blackwell, pp. 283-312.

McDowell, John

[1984] 'Wittgenstein on Following a Rule', *Synthese*, 58, pp. 325-363.

[1981] 'Noncognitivism and Rule-Following' in Holtzman, Steven and Leich,  
Christopher (eds) *Wittgenstein: To Follow a Rule*, London: Routledge and  
Kegan Paul, pp. 141-62.

McGinn, Colin

[1984] *Wittgenstein on Meaning*, Oxford: Blackwell.

McGuinness, Brian

[1982] (ed) *Wittgenstein and his Times*, England: Thoemmes Press.

McManus, Denis

- [1995] 'The Epistemology of Self-Knowledge and the Presuppositions of Rule-Following', *The Monist*, 78(4), pp. 496-514

Miller, Alexander

- [2004] 'Rule Following and Externalism', *Philosophy and Phenomenological Research*, LXVIII(1).
- [2000] 'Horwich, Meaning, and Kripke's Wittgenstein', *Philosophical Quarterly*, 50(199), pp. 161-174.
- [1998] [a] *Philosophy of Language*, McGill-Queen's University Press. [Chapters 5 & 6]
- [1998] [b] 'Rule Following, Response Dependence, and McDowell's Debate with Anti-Realism', *European Review of Philosophy*, 3, pp. 175-97.
- [1997] 'Boghossian on Reductive Dispositionalism about Content: The Case Strengthened' *Mind and Language*, 12(1), pp. 1-10.
- [1989] 'An Objection to Wright's Treatment of Intention', *Analysis*, 49, pp. 169-73.

Miller, Alexander & Divers, John

- [1994] 'Best Opinion, Intention-Detecting, and Analytical Functionalism', *Philosophical Quarterly*, 44, pp. 239-45.

Millikan, Ruth

- [1990] 'Truth Rules, Hoverflies, and the Kripke-Wittgenstein Paradox', *The Philosophical Review*, XCIX(3).

Monk, Ray

- [1990] *Ludwig Wittgenstein: The Duty of Genius*, New York: The Free Press.

Moore, G. E.

[1959] *Philosophical Papers*, London: Allen and Unwin.

[1955] [a] 'Wittgenstein's Lectures in 1930-33, Part III', *Mind*, 64, pp. 1-27.

[1954] [b] 'Wittgenstein's Lectures in 1930-33, Part I', *Mind*, 63, pp. 1-15.

[1954] 'Wittgenstein's Lectures in 1930-33, Part II', *Mind*, 63, pp. 289-315.

Neta, Ram

[1997] 'How Can There be Semantic Facts?', *Southwest Philosophy Review*.

Orwell, George

[1981] *A Collection of Essays*, Harcourt Brace and Company.

Peacocke, Christopher

[1984] 'Review of Wittgenstein on Rules and Private Language', *Philosophical Review*, XCIII, 2.

Pettit, Phillip

[1990] 'The Reality of Rule-Following', *Mind*, 99.

Pettit, Phillip & McDowell, John

[1986] *Subject, Thought and Context*, Oxford: Oxford University Press.

Puhl, Klaus

[1991] *Meaning Skepticism*, New York: de Gruyter.

Putnam, Hilary

[1975] [a] 'The Meaning of Meaning' in Putnam [1975b], pp. 215-72.

[1975] [b] *Mind, Language, and Reality: Philosophical Papers 2*, Cambridge: Cambridge University Press.

[1970] 'Is Semantics Possible?', in Putnam [1975b], pp. 139-53.

Rhees, R.

[1984] (ed) *Recollections of Wittgenstein*, Oxford.

Ryle, Gilbert

[1984] *The Concept of Mind*, Chicago University Press.

Sartorelli, Joseph

[1991] 'McGinn on Concept Skepticism and Kripke's Skeptical Argument', *Analysis*,  
pp. 79-84.

Sartre, Jean Paul

[1956] *Being and Nothingness*, Washington Square Press.

Schiffer, Stephen

[1986] 'Kripkenstein Meets the Remnants of Meaning', *Philosophical Studies*, 49,  
pp. 146-162.

Scruton, Roger

[1984] 'Critical Notice: *Wittgenstein on Rules and Private Language*', *Mind*, XCIII,  
pp. 592-602.

Soames, Scott

[2003] *Philosophical Analysis in the Twentieth Century, v. II: The Age of Meaning*,  
New Jersey: Princeton University Press. [Chapter 2]

[1998] [a] 'Facts, Truth Conditions, and the Skeptical Solution to the Skeptical  
Paradox' in James Tomberlin (ed) *Mind, Language and Ontology*,  
Philosophical Perspectives, 12, Cambridge: Blackwell.

[1998] [b] *Understanding Truth*, Oxford: Blackwell. [Chapter 8]

[1997] 'Skepticism about Meaning: Indeterminacy, Normativity, and the Rule Following Paradox' in *Meaning, Reference, and Truth, The Canadian Journal of Philosophy, Supplementary Volume*.

Strawson, P. F.

[1954] 'Critical Study of the *Philosophical Investigations*', *Mind*, 63, pp. 70-99.

Thornton, Tim

[1997] 'Intention, Rule Following and the Strategic Role of Wright's Order of Determination Test', *Philosophical Investigations* 20(2).

Travis, Charles

[1986] (ed) *Meaning and Interpretation*, Oxford: Blackwell.

Vasiliou, Iakovos

[2001] 'Wittgenstein, Religious Belief, and *On Certainty*' in Robert Arrington and M. Addis (eds), *Wittgenstein and Philosophy of Religion*, pp. 29-50.

Weir, Alan & Miller, Alexander

[2003] 'Objective Content' *Aristotelian Society*, v. 77, pp. 73-90.

Weitz, Morris

[1956] 'The Role of Theory in Aesthetics', *The Journal of Aesthetics and Art Criticism* 15. Page references are to Steve Cahn (ed) *Philosophy for the 21<sup>st</sup> Century*, Oxford: 2003.

Wilson, George

[1998] 'Semantic Realism and Kripke's Wittgenstein', *Philosophy and Phenomenological Research* 58(1).

[1994] 'Kripke on Wittgenstein and Normativity', *Midwest Studies in Philosophy*,  
XIX.

Wittgenstein, Ludwig

[1997] *The Big Typescript*, New York: Oxford Press. (Abbreviated as MS)

[1994] (first published 1956) *Remarks on the Foundations of Mathematics*, G. E. M. Anscombe (trans), G. H. von Wright, R. Rhees, and G. E. M. Anscombe (ed), MIT Press. (Abbreviated as RFM)

[1982] D. Lee (ed), *Wittgenstein's Lectures: Cambridge, 1930-32*, Chicago: University of Chicago Press. (Abbreviated as WL)

[1980] [a] G. H. von Wright (ed), P. Winch (trans), *Culture and Value*, Chicago: University of Chicago Press. (Abbreviated as CV)

[1980] [b] *Remarks on the Philosophy of Psychology*, G. E. M. Anscombe (trans), G. H. von Wright and G. E. M. Anscombe (ed), Oxford: Blackwell. (Abbreviated as RPP)

[1976] Cora Diamond (ed), *Wittgenstein's Lectures on the Foundations of Mathematics: Cambridge 1939*, Ithaca: Cornell University Press. (Abbreviated as LFM)

[1975] R. Rhees (ed), R. Hargreaves and R. White (trans), *Philosophical Remarks*, Oxford: Basil Blackwell. (Abbreviated as PR)

[1974] Anthony Kenny (trans), R. Rhees (ed), *Philosophical Grammar*, Oxford: Blackwell. (Abbreviated as PG)

[1961] G. E. M. Anscombe (ed and trans), *Notebooks, 1914-1916*, Oxford: Basil Blackwell (Abbreviated as NB)

- [1960] (first published 1922) D. Pears and B. F. McGuinness (trans), *Tractatus Logico Philosophicus*, Oxford: Blackwell. (Abbreviated as TLP)
- [1958] *The Blue and Brown Books*, New York: Harper and Row. (Abbreviated as BB)
- [1953] G. E. M. Anscombe (trans), G. E. M. Anscombe and R. Rhees (eds), *Philosophical Investigations*, Oxford: Blackwell. (Abbreviated as PI)

Wright, Crispin

- [2007] 'Rule-Following without Reasons: Wittgenstein's Quietism and the Constitutive Question' in John Preston (ed); *Ratio* XX(4). [Page reference are to the PDF online version of this article.]
- [2001] *Rails to Infinity: Essays on Themes from Wittgenstein's Philosophical Investigations*, Harvard University Press.
- [1999] 'Truth: A Traditional Debate Reviewed' in Simon Blackburn and Keith Simmons (eds) *Truth*, Oxford: Oxford Press.
- [1998] 'Comrades Against Quietism', *Mind*, 107.
- [1992] *Truth and Objectivity*, Cambridge Mass: Harvard University Press.
- [1987] 'Realism, Antirealism, Irrealism, Quasi-Realism', *Midwest Studies in Philosophy* XII, pp. 29-47.
- [1986] *Realism, Meaning and Truth*, Oxford: Blackwell.
- [1980] *Wittgenstein on the Foundations of Mathematics*, Harvard. [Part I]

Yalowitz, Steven

- [2000] 'A Dispositional Account of Self-Knowledge', *Philosophy and Phenomenological Research*, LXI(2).

Zalabardo, Jose

[1997] 'Kripke's Normativity Argument', *Canadian Journal of Philosophy*, 27(4),  
pp. 467-488.

### *Autobiographical Statement*

Chrysoula Gitsoulis was born in Thessaloniki, Greece, and raised in Asbury Park, on the Jersey Shore, back when it was the Music Mecca of the East Coast. She received a B.A. in Philosophy, B.S. in Mathematics, and Minor in Classics from Rutgers College, New Jersey, in 1991. Upon graduation, she served for four years as a part time tutor and teaching assistant for the Math Division of the *Educational Opportunity Fund Program* at Douglass College, Rutgers University. After that, she enrolled in the Ph.D Program in Philosophy at the Graduate Center of the City University of New York. She is currently an Adjunct Lecturer in Philosophy at City College of the City University of New York, and has taught so many classes as a graduate student that she believes she holds the department record! (Or, if not that one, the record on # of dissertation revisions!) Her research interests focus on the History of Philosophy (especially Plato, Aristotle, and Wittgenstein), and Contemporary Moral and Social Issues. She has a passion for music (esp. Rock & Blues), and greatly enjoys travel (esp. to unspoiled landscapes and secluded, traditional villages), hiking, art-making, photography, the Jersey Shore, and the Aegean Sea.